Conduction band nonparabolicity, chemical potential, and carrier concentration of intrinsic InSb as a function of temperature

Stefan Zollner,[1, a)] Carlos A. Armenta,[1] Sonam Yadav,[1] and José Menéndez[2]
[1)] *Department of Physics, New Mexico State University, MSC 3D, PO Box 30001, Las Cruces, NM 88003-8001, USA*
[2)] *Department of Physics, Arizona State University, Tempe, AZ, 85287-1504, USA*

(Dated: 29 December 2023)

In this review, the nonparabolicity of the light hole and electron bands at the Γ-point in cubic diamond or zinc blende semiconductors is derived from Kane's 8×8 $\vec{k}\cdot\vec{p}$ model in the large spin-orbit splitting approximation. Examples of several approximations are given with InSb as an example and their accuracy is discussed. To determine the temperature dependence of the effective masses and the nonparabolicity parameters, the unrenormalized band gap must be utilized. This includes only the redshift of the band gap due to thermal expansion, not the renormalization due to deformation potential electron-phonon coupling. As an application of this method, the chemical potential and the charge carrier concentration of intrinsic InSb are calculated from 50 to 800 K and compared with electrical and optical experiments. These results are also relevant for other semiconductors with small band gaps as needed for mid-infrared detector applications.

I. INTRODUCTION

The curvature at the bottom of the lowest conduction band (CB) in cubic zinc blende semiconductors, such as InSb or GaAs, determines many processes, including electron transport, low-temperature specific heat, and the absorption and emission of light.[1] For bands with spherical symmetry, especially at the Γ-point, it can be expressed as a series of even powers of the wave vector k, because terms with odd powers are small.[2,3] In the parabolic band approximation, which is treated in many textbooks[1,4] and often sufficient, the unrenormalized CB energy is written as

$$E_e^u(k) = E_0^u + \frac{\hbar^2 k^2}{2m_0}\frac{1}{m^*}, \tag{1}$$

where E_0^u is the unrenormalized direct band gap (we will explain later what that means), $\hbar$ the reduced Planck's constant, m_0 the free electron mass, and the dimensionless parameter m^* the effective electron mass. This parabolic expression (1) is valid only if the second term is much smaller than the band gap E_0^u. Especially in semiconductors with small band gaps such as InSb higher-order terms must also be considered. That is the topic of this review.

Sophisticated *ab initio* band structure calculations are available to obtain many properties of semiconductors, but they lack transferability. One calculation is only valid for a single material. Changing the composition of a semiconductor alloy or selecting a different compound will usually require a new calculation. This limits the predictive power of *ab initio* calculations. The goal of our work is transferability and simplicity resulting from analytical expressions. Starting from Kane's 8×8 $\vec{k}\cdot\vec{p}$-model,[1,2,5] we will show that only a small number of parameters, especially the band gap and one momentum matrix element,[6] are sufficient to predict many semiconductor properties related to the CB nonparabolicity. While we will focus our discussion on InSb, the transferability of our model allows applications to other infrared detector materials, such as InAs and alloys like SiGeSn, InGaAsSb, or HgCdTe.

Our starting point is the classical 1957 paper by Kane on the *Band structure of indium antimonide*.[2] We simplify Kane's model and only include its essential elements to allow analytical treatment of the results. We bring this model up to date with current experimental results, especially regarding the temperature dependence and the renormalization of band energies due to the deformation-potential electron-phonon interaction. A recent treatment of the CB nonparabolicity was also presented by Masut.[7] Our work is similar in some aspects, but we avoid the introduction of triple-index generalized Fermi-Dirac integrals.[8–10] Instead, we use Fermi-Dirac integrals $F_n(x)$ that can be evaluated in MATLAB[11] using polylogarithm functions.[12] We discuss the validity of our approximations, present graphical representations of our results, and include our MATLAB scripts and detailed derivations as supplementary material.[13] We apply our nonparabolicity model to calculate the chemical potential and the free carrier concentration of intrinsic InSb as a function of temperature and compare with experimental results.

II. THEORETICAL MODEL

A. Notation and conventions

We begin by introducing some symbols and notation to allow compact expressions for the electronic band structure. $E_n(k)$ is the energy of a band as a function of wave

[a)] zollner@nmsu.edu; http://femto.nmsu.edu

vector k. This energy is positive in the CB and negative in the valence band (VB). The subscript n is the band index for the conduction band (e) or the split-off, light, and heavy hole bands (so, lh, hh). $\epsilon_n(k)$ is the energy above or below the band extremum. This is always positive. We use superscripts to distinguish between the experimental (exp) and unrenormalized (u) band energies.

Expressions of band energies resulting from $\vec{k}\cdot\vec{p}$-theory can be simplified, if the kinetic energy of the free electron is subtracted from the band energies. Kane[2,14] therefore introduced a modified energy parameter

$$\tilde{E}_n^u\left(\vec{k}\right) = E_n^u\left(\vec{k}\right) - \frac{\hbar^2 k^2}{2m_0}. \tag{2}$$

We use a tilde instead of a prime in Eq. (2), because the prime (as in E_0', for example) has taken a different meaning in recent years. It denotes optical interband transitions, also known as critical points, into the p-antibonding conduction band.

B. Kane's $\vec{k}\cdot\vec{p}$ model and solution for large SO splitting

An electronic band structure method called $\vec{k}\cdot\vec{p}$-theory is based[15] on the Bloch wave form of the wave function $\psi_{n\vec{k}}(\vec{r}) = u_{n\vec{k}}(\vec{r})\exp(i\vec{k}\cdot\vec{r})$, where $u_{n\vec{k}}(\vec{r})$ is periodic in the crystal lattice.[1] We assume that the solution of the time-independent Schrödinger equation $\tilde{H}u_{n0} = \tilde{E}_{n0}^u u_{n0}$ is known at the at the Γ-point for $\vec{k}$=0 with wave functions u_{n0} and eigenvalues $\tilde{E}_{n0}^u$, for example from experimental measurements of the band energies E_{n0}^u. $\tilde{H}$ is the Hamiltonian where the free electron kinetic energy $\hbar^2k^2/2m_0$ has been subtracted. The energies and wave functions for small nearby $\vec{k}$ can then be obtained in perturbation theory by solving the eigenvalues of the matrix[1]

$$\sum_i \left(E_{n0}^u\delta_{ni} + \frac{\hbar}{m_0}\vec{k}\cdot\langle n0|\,\vec{p}\,|i0\rangle\right)c_{ni} = \tilde{E}_{n\vec{k}}^u c_{nn}. \tag{3}$$

$\langle n0|\,\vec{p}\,|i0\rangle$ is the momentum matrix element connecting the bands with indices n and i at the Γ-point, which is also known as the $\vec{k}\cdot\vec{p}$ matrix element, and related to the optical dipole matrix element.[4] Details of this method are included in many textbooks,[1] review articles,[5] and in the supplementary material.[13]

For practical purposes, one starts with deciding how many bands should be included in the calculation. This determines the dimension of the eigenvalue problem given by Eq. (3). For this work, we only include the three top VBs (the p-bonding bands) and the lowest CB (the s-antibonding band). At the Γ-point, we select wave functions $|S\rangle$ for the CB and $|X \pm iY\rangle$, $|Z\rangle$ for the VB. Without loss of generality, we may assume that $\vec{k}$ points along the z-direction. The only non-vanishing momentum matrix elements are of the form $\langle S|\,p_x\,|X\rangle = -iP$. The mixed momentum matrix elements $\langle S|\,p_x\,|Y\rangle$ etc. vanish. Counting the spin degeneracy, this yields an 8×8 matrix, with two identical 4×4 on-diagonal block matrices[2,5]

$$\tilde{H}_{\vec{k}} = \begin{pmatrix} E_0^u & 0 & -\frac{\hbar k}{m_0}iP & 0 \\ 0 & -\frac{2\Delta_0}{3} & \frac{\sqrt{2}\Delta_0}{3} & 0 \\ \frac{\hbar k}{m_0}iP & \frac{\sqrt{2}\Delta_0}{3} & -\frac{\Delta_0}{3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{4}$$

and vanishing off-diagonal blocks. Δ_0 is the matrix element of the spin-orbit (SO) Hamiltonian, also known as the SO splitting. To simplify the notation, one introduces the energy $E_P = 2P^2/m_0$, which has values between 18 and 26 eV for many semiconductors.[6] More sophisticated $\vec{k}\cdot\vec{p}$-models will include more bands, which requires the knowledge of other energy gaps and additional matrix elements. For example, one might include all s- and p-bonding and antibonding bands (which leads to a 16×16 matrix) or bands with d-type symmetry (30×30).[17]

The matrix (4) has one obvious eigenvalue $\tilde{E}^u$=0. This solution is identified with the heavy hole band. Its energy has the wrong sign and is equal to the kinetic energy of the free electron. The downward curvature and warping of this heavy hole band are caused by higher-lying CBs,[14,16] which we have neglected in our simple model. We do not consider this solution for our review and instead use the experimental parabolic density-of-states heavy hole mass m_{hh}=0.43 determined from Hall effect measurements for our calculations.[18]

The other three eigenvalues of the matrix (4) are determined from the cubic characteristic equation[2]

$$\tilde{E}^u\left(\tilde{E}^u - E_0^u\right)\left(\tilde{E}^u + \Delta_0\right) - \frac{\hbar^2k^2E_P}{2m_0}\left(\tilde{E}^u + \frac{2\Delta_0}{3}\right) = 0, \tag{5}$$

which can be solved analytically as described in the supplementary material[13] and shown in Fig. 1. For our purposes, these analytical solutions to the cubic equation are not useful, because they cannot be inverted to yield the density of states as a function of excess energy ϵ_n.

For very small values of k, the characteristic equation (5) can be solved perturbatively, leading to the effective masses of the electron, split-off and light hole bands[2,5]

$$\frac{1}{m_{\rm lh}^*} = \frac{2}{3}\frac{E_P}{E_0^u} - 1, \tag{6}$$

$$\frac{1}{m_{\rm so}^*} = \frac{E_P}{3\left(E_0^u + \Delta_0\right)} - 1, \tag{7}$$

$$\frac{1}{m_e^*} = 1 + \frac{E_P}{3}\left(\frac{2}{E_0^u} + \frac{1}{E_0^u + \Delta_0}\right), \tag{8}$$

as shown in the supplementary material.[13] Due to the large nonparabolicity of the bands, these effective masses can only be applied for very small values of k, see the dotted lines in Fig. 1.

To obtain a simple analytical solution of Eq. (5), we use the large SO splitting approximation $\tilde{E}^u \ll \Delta_0$. The characteristic equation (5) then becomes quadratic and

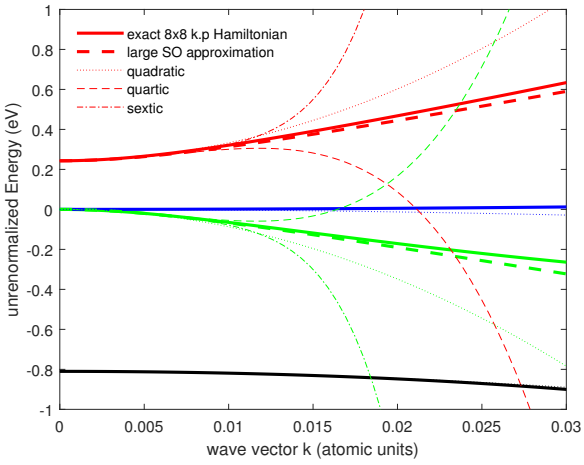

FIG. 1. Band structure for InSb at 0 K. Thick lines show the heavy hole (blue), light hole (green), split-off hole (black), and electron bands (red) from the cubic characteristic equation (5) (solid) and from the large SO approximation (9) (dashed) as a function of the wave vector k in atomic units (inverse Bohr radii). Thin lines show the expansion of the square root in Eq. (9) including terms proportional to k^2 (dotted), k^4 (dashed), and k^6 (dot-dashed) for the electron and light hole bands. Parabolic bands for the heavy and split off holes with experimental masses are also shown (dotted). See supplementary material[13] for a similar graph showing the energies as a function of the square of wave vector.

offers solutions for the light hole and electron bands[2,5]

$$E^u_{e,lh} = \frac{\hbar^2 k^2}{2m_0} + \frac{E^u_0}{2}\left(1 \pm \sqrt{1 + \frac{\hbar^2 k^2}{2m_0}\frac{2}{\mu_{lh}E^u_0}}\right) \quad (9)$$

with effective and reduced masses

$$m^*_e = \frac{3E^u_0}{2E_P + 3E^u_0}, \quad (10)$$

$$m^*_{lh} = \frac{3E^u_0}{2E_P - 3E^u_0}, \quad (11)$$

$$\mu_{lh} = \frac{m^*_e m^*_{lh}}{m^*_e + m^*_{lh}} = \frac{3E^u_0}{4E_P} \quad (12)$$

obtained by keeping only the lowest order terms in Eq. (9). Since the matrix element E_P is much larger than the band gap E^u_0, the light hole and electron masses are nearly the same. The square root in Eq. (9) can be expanded into powers of k^2. Unfortunately, this series only converges for small values of k, see Fig. 1. The large SO approximation is very good and Eq. (9) represents the electron and light hole solutions of the characteristic Eq. (5) quite well, see the small difference between the solid and dashed lines in Fig. 1.

C. Nonparabolicity parameters

To calculate the density of states, which is important for thermal and transport properties, we need to invert Eq. (9) and write k^2 as a function of energy. This requires solving a quadratic equation, which results in[19,20]

$$\frac{\hbar^2 k^2}{2m_0} = \epsilon^u_e + \frac{E^u_0}{2m^*_e}\left(1 - \sqrt{1 + \frac{2\epsilon^u_e m^{*2}_e}{\mu_{lh}E^u_0}}\right) \quad \text{and} \quad (13)$$

$$\frac{\hbar^2 k^2}{2m_0} = E^u_{lh} + \frac{E^u_0}{2m^*_{lh}}\left[1 - \sqrt{1 + \frac{2E^u_{lh}m^{*2}_{lh}}{\mu_{lh}E^u_0}}\right] \quad (14)$$

for the CB and light hole band, respectively.

The nonparabolicity coefficients α_n and β_n are defined by[19,20]

$$\frac{\hbar^2 k^2}{2m_0 m_n} = \epsilon^u_n\left(1 + \alpha_n\epsilon^u_n + \beta_n\epsilon^{u2}_n\right). \quad (15)$$

They can be obtained[19] by expanding Eqs. (13,14) into a power series of ϵ_n

$$\frac{\hbar^2 k^2}{2m_0 m_e \epsilon^u_e} = 1 + \frac{m^{*2}_e\epsilon^u_e}{4\mu^2_{lh}E^u_0} - \frac{m^{*4}_e\epsilon^{u2}_e}{4\mu^3_{lh}E^{u2}_0} \quad \text{and} \quad (16)$$

$$\frac{\hbar^2 k^2}{2m_0 m_{lh}\epsilon^u_{lh}} = 1 + \frac{m^{*2}_{lh}\epsilon^u_{lh}}{4\mu^2_{lh}E^u_0} + \frac{m^{*4}_{lh}\epsilon^{u2}_{lh}}{4\mu^3_{lh}E^{u2}_0}. \quad (17)$$

The nonparabolicity parameters in the large SO splitting approximation from an 8×8 $\vec{k}\cdot\vec{p}$-model are therefore[19–21]

$$\alpha_e = \frac{m^{*2}_e}{4\mu^2_{lh}E^u_0} = \frac{(1-m_e)^2}{E^u_0}, \quad (18)$$

$$\beta_e = -\frac{m^{*4}_e}{4\mu^3_{lh}E^{u2}_0} = \frac{-2m^*_e(1-m^*_e)^3}{E^{u2}_0}, \quad (19)$$

$$\alpha_{lh} = \frac{m^{*2}_{lh}}{4\mu^2_{lh}E^u_0} = \frac{(1+m_{lh})^2}{E^u_0}, \quad (20)$$

$$\beta_{lh} = \frac{m^{*4}_{lh}}{4\mu^3_{lh}E^{u2}_0} = \frac{2m^*_{lh}(1+m^*_{lh})^3}{E^{u2}_0}. \quad (21)$$

Figure 2 compares the "exact" solution of the 8×8 $\vec{k}\cdot\vec{p}$-Hamiltonian in the large SO splitting approximation given by Eqs. (13,14) for the light hole and electron bands with those obtained by expansion with the nonparabolic corrections (18-21). The error obtained with just the lowest-order nonparabolicity correction (β_n=0) is about the same as the error caused by the large SO approximation. Adding the next term ($\beta_n\neq$0) makes the solution nearly indistinguishable from the exact large SO-approximation. For this review, we therefore set β_n=0 for applications of our nonparabolic band structure model.

D. Temperature dependence of the effective masses

Previously, we referred to E^u_0 as the unrenormalized band gap, but we did not explain what that means.

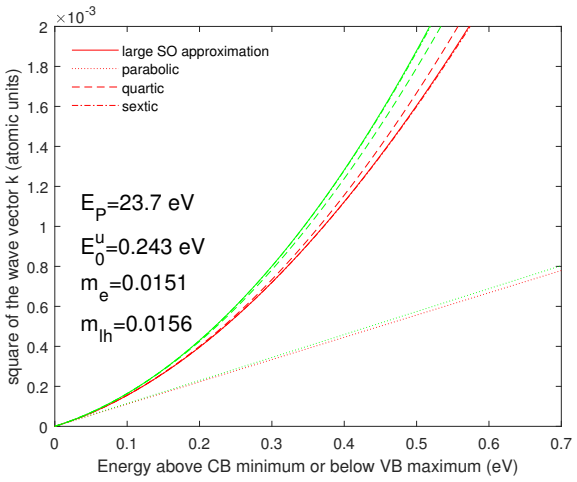

FIG. 2. The square of the wave vector (in atomic units) versus energy from the extremum for electrons (red) and light holes (green) in InSb at 0 K, calculated using Eqs. (13,14) within the large SO splitting approximation (solid), using parameters from Sec. II E. The dotted lines show the parabolic band dispersion. The dashed and dot-dashed lines show the next two terms in the Taylor expansions of the square root. Only the first term (β_n=0) gives a good approximation (dashed). If the β_n-term is included (dot-dashed), the deviation from the exact square-root expressions (13,14) is nearly indistinguishable on this scale.

The effective electron and light hole masses (10,11) depend on temperature, because E_0^u varies with temperature. (We ignore the small variation of the momentum matrix element E_P due to thermal expansion.[22]) However, the matter is complicated. One **cannot** simply use the experimental band gap $E_0^{\rm exp}$ to calculate the effective masses.[7,23] ($E_0^{\rm exp}$ is the energy separation between the bottom of the CB and the top of the VB. It is also called the "thermal gap", because it enters the calculation of the carrier concentration using Fermi-Dirac statistics. The onset of optical absorption, aka the optical activation energy, may be larger than the band gap due to the Burstein-Moss shift.[24,25] We do not consider band gap renormalization due to many-body effects in this work, although it might play a role at higher temperatures and the corresponding high intrinsic carrier concentrations.[26,27])

Instead, we look into the origins of the temperature dependence of the band gap and its three contributions:[28,29]

$$\frac{\partial E_0^{\rm exp}(T)}{\partial T} = \left(\frac{\partial E_0}{\partial T}\right)_{\rm TE} + \left(\frac{\partial E_0}{\partial T}\right)_{\rm DW} + \left(\frac{\partial E_0}{\partial T}\right)_{\rm SE}. \quad (22)$$

The first term[30]

$$\left(\frac{\partial E_0}{\partial T}\right)_{\rm TE} = -3\alpha(T)\, B \left(\frac{\partial E_0}{\partial p}\right)_T, \quad (23)$$

describes the redshift of the band gap due to thermal expansion. It can be calculated from the temperature-dependent thermal expansion coefficients $\alpha(T)$,[31–34] the temperature-independent bulk modulus B=46 GPa (from Ref. 35), and the pressure coefficient of the band gap at constant temperature[36] equal to 0.155 eV/GPa.

The second and third terms in Eq. (22) are the Debye-Waller and self-energy (SE) terms, respectively, which describe the renormalization of the band gap due to deformation-potential electron-phonon interaction. The Debye-Waller term arises from the second-order electron-phonon Hamiltonian (simultaneous absorption or emission of two phonons by an electron) taken to first order in perturbation theory and is usually negative.[29] The self-energy term arises from the first-order electron-phonon Hamiltonian taken to second order in perturbation theory (emission/absorption of a phonon by an electron followed by reabsorption/reemission) and is often positive.[29] The theory of this electron-phonon renormalization of the band gap has been described by Cardona and Gopalan.[28] An application of this theory to the direct band gap of InSb was given in Ref. 29. Because of the zero-point phonon oscillations, the electron-phonon interaction renormalizes the band gap even at 0 K. Therefore, E_0^u and $E_0^{\rm exp}$ are different at 0 K. The latter includes the electron-phonon contribution (renormalization), while the former does not.

The temperature dependence of the band gap due to thermal expansion from Eq. (23) is

$$E_0^u(T) = E_0^u(0\ {\rm K}) - 3B\left(\frac{\partial E_0}{\partial p}\right)_T \int_0^T \alpha(\theta)\, d\theta. \quad (24)$$

We call this the "mass band gap" or the unrenormalized temperature-dependent band gap, because the renormalization of the band gap due to deformation potential electron-phonon interaction given by the last two terms in Eq. (22) has not been included. **The energy in Eq. (24) must be used to calculate the temperature-dependent effective masses and band structures from $\vec{k}\cdot\vec{p}$-theory.**

The thermal expansion coefficient of zinc blende semiconductors is[37]

$$\alpha(T) = A\left(\frac{T}{\Theta_D}\right)^3 I_D\left(\frac{\Theta_D}{T}\right), \quad (25)$$

where

$$I_D(x_D) = \int_0^{x_D} \frac{x^4 e^x dx}{(e^x-1)^2} \quad (26)$$

is the Debye integral (which can be solved numerically in MATLAB[11]), A is an adjustable parameter, and Θ_D the Debye temperature. The Debye temperature for InSb is about 168 K for InSb,[32] but we treat it as an adjustable parameter to fit the thermal expansion coefficients.[31] With parameters A=17.5×10^{-6} K^{-1} and Θ_D=450 K, satisfactory agreement with the experimental data can be achieved above 100 K, see Fig. 3.

The agreement can be improved, especially at low temperatures, by separately considering the contributions of

transverse acoustic (TA), longitudinal acoustic (LA), and optical (O) phonons[37]

$$\alpha(T) = -A_{\rm TA}\left(\frac{\Theta_{\rm TA}}{T}\right)^2 \frac{\exp(\Theta_{\rm TA}/T)}{[\exp(\Theta_{\rm TA}/T)-1]^2} + \\ + A_{\rm LA}\left(\frac{T}{\Theta_{\rm LA}}\right)^3 I_D\left(\frac{\Theta_{\rm LA}}{T}\right) + \\ + A_O\left(\frac{\Theta_O}{T}\right)^2 \frac{\exp(\Theta_O/T)}{[\exp(\Theta_O/T)-1]^2}. \quad (27)$$

The vibrational properties of InSb lead to phonon parameters $\Theta_{\rm TA}$=53.7 K, $\Theta_{\rm LA}$=314 K, and Θ_O=240 K calculated as described in Ref. 37. The amplitudes are obtained as fit parameters: $A_{\rm TA}=3\times10^{-6}$, $A_{\rm LA}=0.11\times10^{-6}$, and $A_O=1.5\times10^{-6}$. Both expressions (25) and (27) lead to the same thermal expansion shift of the band gap (24), i.e., the negative thermal expansion coefficient at low temperatures is not a large contribution.

To determine E_0^u (0 K), we proceed as follows: The experimental band gap was determined to be[38]

$$E_0^{\rm exp}(T) = E_B^{\rm exp} - a_B^{\rm exp}\left[1+\frac{2}{\exp(\Omega/k_BT)-1}\right] \quad (28)$$

with parameters $E_B^{\rm exp}$=261 meV (unrenormalized band gap), $a_B^{\rm exp}$=26 meV (electron-phonon coupling strength), and Ω=18.9 meV (energy of the coupling phonon, correcting an error in Ref. 38). This result (28) overestimates the electron-phonon parameters, because it includes the redshift due to thermal expansion as well as due to the renormalization by electron-phonon interactions. To calculate the pure electron-phonon term

$$E_0^{\rm el-ph}(T) = E_0^{\rm exp}(0\ {\rm K}) - 3B\left(\frac{\partial E_0}{\partial p}\right)_T \int_0^T \alpha(\theta)\, d\theta, \quad (29)$$

we subtract the thermal expansion contribution from the experimental band gap and fit the difference (29) with a Bose-Einstein expression as given in Eq. (28). This results in parameters $E_B^{\rm el-ph}$=243 meV, $a_B^{\rm el-ph}$=7.3 meV, and $\Omega^{\rm el-ph}$=10.9 meV. By definition, the unrenormalized band gap E_0^u (0 K) is equal to $E_B^{\rm el-ph}$. We are now able to calculate $E_0^u(T)$ with Eq. (24), which determines the $\vec{k}\cdot\vec{p}$-band structure.

The experimental band gap $E_0^{\rm exp}(T)$ and the contributions due to thermal expansion $E_0^{\rm TE}(T)$ and electron-phonon renormalization $E_0^{\rm el-ph}(T)$ are shown in Fig. 4. The latter two are similar in magnitude over the complete range. About half of the redshift of the direct band gap with increasing temperature is caused by thermal expansion, the other half by deformation-potential electron-phonon interactions. The unrenormalized temperature-dependent band gap $E_0^u(T)$ follows the thermal expansion contribution $E_0^{\rm TE}(T)$, but is shifted upward by 8 meV due to the renormalization of the low-temperature band gap by zero-point phonon vibrations. (The electron-phonon shift obtained with the rigid pseudo-ion method was larger.[29]) We are now able to calculate the temperature dependence of the effective electron and light hole masses using Eqs. (10,11). Results are shown in Fig. 5.

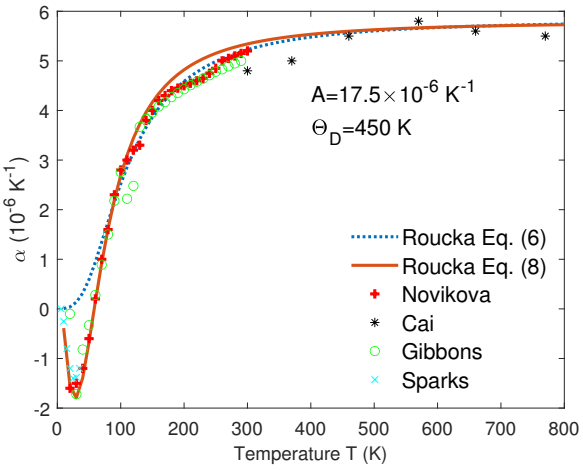

FIG. 3. Linear thermal expansion coefficient α versus temperature taken from the literature[31–34] (symbols) along with a fit to the data using Eq. (25) (solid).

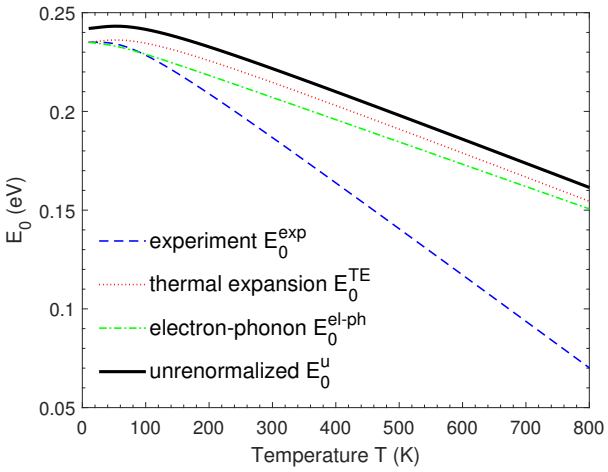

FIG. 4. Direct gap E_0 of InSb versus temperature. Dashed: experimental gap $E_0^{\rm exp}$ from Eq. (28); dotted: thermal expansion contribution; dot-dashed: electron-phonon contribution from Eq. (29); solid: unrenormalized band gap from Eq. (24) for calculation of effective masses.

E. Momentum matrix element and effective masses

It has been standard practice[7,23] to calculate $\vec{k}\cdot\vec{p}$-band structures from the temperature-dependent band gap $E_0^{\rm TE}$ that includes thermal expansion, but neglects the renormalization due to electron-phonon coupling at ele-

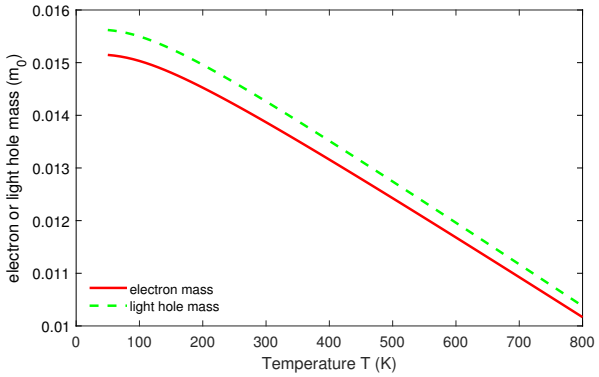

FIG. 5. Effective masses of the electron (solid) and light hole (dashed) bands of InSb as a function of temperature, calculated taking into account only the thermal expansion contribution to the band gap, not the electron-phonon renormalization.

vated temperatures. In other words, one first calculates the unrenormalized band structure in $\vec{k}\cdot\vec{p}$-perturbation theory and then adds the electron-phonon coupling as a second perturbation. This approach is inconsistent, however, because the renormalization due to the zero-point phonon vibration is included in $E_0^{\rm TE}$, while renormalization due to thermal excitation of phonons is not. We prefer an approach where the entire electron-phonon coupling is treated as a whole, given by the square brackets in Eq. (28). Therefore, the energy $E_0^u(T)$ should enter the $\vec{k}\cdot\vec{p}$-band structure.

This standard practice has resulted in a comprehensive body of work, especially the compilation of matrix element parameters by Lawaetz for many different semiconductors.[6] Lawaetz calculated $\vec{k}\cdot\vec{p}$-parameters based on cyclotron measurements of the effective masses and experimental low-temperature band gaps, which include the renormalization due to zero-point phonon vibrations. To exclude *all* electron-phonon coupling effects consistently, we need to fine-tune the momentum matrix elements, especially for semiconductors with small band gaps used for mid-infrared optical detector applications.

Starting with the cyclotron light hole mass m_{lh}^*=0.0156 for InSb at low temperature and the unrenormalized band gap E_0^u=0.243 eV, we solve Eq. (6) to obtain

$$E_P = \frac{3}{2}E_0^u\left(1+\frac{1}{m_{lh}^*}\right) = 23.7 \text{ eV}, \tag{30}$$

somewhat larger than the usual value of 23.1 eV published by Lawaetz.[6] The corresponding effective electron mass (8) with Δ_0=0.81 eV equals m_e^*=0.0136, which is in excellent agreement with the experimental value. If we instead use the expressions (10,11) from the large SO splitting approximation, then the light hole mass remains the same, but the effective electron mass increases to 0.0151. This is a small price we need to pay for the ease of our analytical approach.

III. APPLICATION TO THERMAL PROPERTIES

A. Density of states

For the calculation of the chemical potential, we need the density of states[39]

$$\begin{aligned} g_n(\epsilon_n) &= \frac{1}{4\pi^3}\int d^3\vec{k}\,\delta\left(E_{n\vec{k}}-\epsilon_n\right) = \\ &= \frac{1}{\pi^2}\int_0^\infty k^2 dk\,\delta\left(E_{nk}-\epsilon_n\right). \end{aligned} \tag{31}$$

We have included the spin degeneracy and assumed that the bands are spherically symmetric.

By taking the derivative of Eq. (15) on both sides, we find[20]

$$dk = \sqrt{\frac{m_0 m_n}{2\hbar^2\epsilon_n}}\frac{1+2\alpha_n\epsilon_n+3\beta_n\epsilon_n^2}{\sqrt{1+\alpha_n\epsilon_n+\beta_n\epsilon_n^2}}d\epsilon_n \quad \text{and} \tag{32}$$

$$\begin{aligned} k^2dk &= \frac{1}{2}\left(\frac{2m_0m_n}{\hbar^2}\right)^{\frac{3}{2}}\times \\ &\times\sqrt{\epsilon_n\left(1+\alpha_n\epsilon_n+\beta_n\epsilon_n^2\right)}\left(1+2\alpha_n\epsilon_n+3\beta_n\epsilon_n^2\right)d\epsilon_n. \end{aligned} \tag{33}$$

The density of states is therefore[7,39]

$$\begin{aligned} g_n(\epsilon_n) &= \frac{1}{2\pi^2}\left(\frac{2m_0m_n^*}{\hbar^2}\right)^{\frac{3}{2}}\times \\ &\times\sqrt{\epsilon_n\left(1+\alpha_n\epsilon_n+\beta_n\epsilon_n^2\right)}\left(1+2\alpha_n\epsilon_n+3\beta_n\epsilon_n^2\right). \end{aligned} \tag{34}$$

We see that the nonparabolicity enhances the density of states by a factor

$$\left(1+2\alpha_n\epsilon_n+3\beta_n\epsilon_n^2\right)\sqrt{1+\alpha_n\epsilon_n+\beta_n\epsilon_n^2} \approx 1+\frac{5}{2}\alpha_n\epsilon_n \tag{35}$$

to first order in $\alpha_n\epsilon_e$ if we set β_n to zero. Since the density of states depends on $m_n^{*\frac{3}{2}}$, we can define an energy-dependent density-of-states effective mass

$$m_{n,\rm DOS}^*(\epsilon_n) = m_n^*\sqrt[3]{1+\alpha_n\epsilon_n+\beta_n\epsilon_n^2}\left(1+2\alpha_n\epsilon_n+3\beta_n\epsilon_n^2\right)^{\frac{2}{3}}. \tag{36}$$

By setting β_e=0 and keeping only terms linear in $\alpha_e\epsilon_e$, the effective electron mass m_e increases approximately like

$$m_{e,\rm DOS}^*(\epsilon) \approx m_e^*\left(1+\frac{5}{3}\alpha_e\epsilon_e\right). \tag{37}$$

In other words, when the excess energy ϵ_e is equal to the band gap E_0 (i.e., $\alpha_e\epsilon_e\approx$1), the effective electron mass m_e^* has nearly tripled. This is shown in Fig. 6, which plots the effective density-of-states electron and light hole masses of InSb as a function of excess energy above the conduction band minimum. Most of the mass enhancement is due to α_e-term (shown by the dotted line). We therefore have confidence that the expansion (31) converges well in the approximation for large SO splittings.

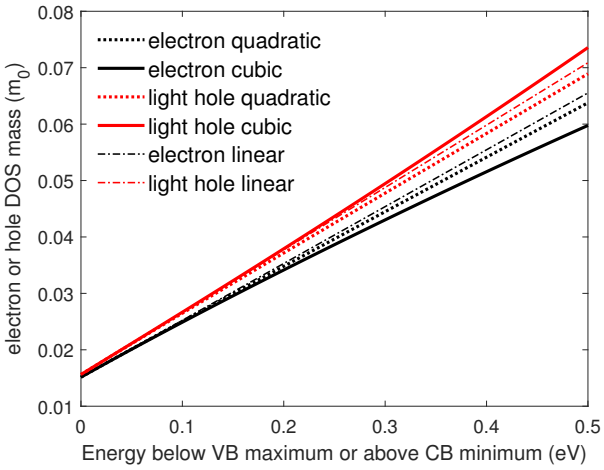

FIG. 6. Effective density of states electron mass $m^*_{e,\mathrm{DOS}}$ (black) and light hole mass $m^*_{lh,\mathrm{DOS}}$ (red) of InSb at 0 K as a function of excess energy above the conduction band minimum or below the valence band maximum, calculated from Eq. (36) (solid). The dotted lines show the results with β_n=0. The dash-dotted lines show the linear expansion (37).

Since $\alpha_e\epsilon_e$ is not exactly small, one might wonder to what extent the linear expansion (37) is accurate. As shown in Fig. 6, the linearization of Eq. (37) introduces a small error, which overestimates the effective electron mass for very high electron energies. This error has a similar magnitude to those introduced by leaving out higher "remote" bands in the $\vec{k}\cdot\vec{p}$-model[2] or with the large SO appxroximation. The linearization (37) of the density of states is necessary to evaluate the chemical potential of a degenerate electron gas using Fermi-Dirac integrals.

B. Chemical potential and intrinsic carrier concentration versus temperature

We apply the density of states (31) for nonparabolic bands to calculate the chemical potential μ and the intrinsic carrier concentration n for InSb as a function of temperature T.

The electron density n_Γ in the Γ-valley of the CB at temperature T is[39,40]

$$n_\Gamma(T) = \int_0^\infty d\epsilon g_e(\epsilon) f\left[E_0^{\mathrm{exp}}(T) + \epsilon_e\right] \quad \text{where} \quad (38)$$

$$f(E) = \left[\exp\left(\frac{E-\mu}{k_BT}\right) + 1\right]^{-1} \quad (39)$$

is the Fermi-Dirac distribution function with the chemical potential μ and the Boltzmann constant k_B. Note that we use the experimental (or "thermal") band gap $E_0^{\mathrm{exp}}(T)$ in Eq. (38), not the "mass" band gap $E_0^u(T)$ introduced in Sec. II D.

By setting β_n=0 and keeping only terms linear in α_n, we find that the density of states enhancement factor (35) is approximately $1+\frac{5}{2}\alpha_e\epsilon_e$. With the substitutions $y=\epsilon_e/k_BT$ and $x=\left(\mu-E_0^{\mathrm{exp}}\right)/k_BT$, the electron density can be written using Fermi-Dirac integrals as[41,42]

$$n_\Gamma(T) = N_e(T)\left[F_{\frac{1}{2}}\left(\frac{\mu-E_0^{\mathrm{exp}}}{k_BT}\right) + \right. \\ \left. + \frac{15}{4}\alpha_e k_BT F_{\frac{3}{2}}\left(\frac{\mu-E_0^{\mathrm{exp}}}{k_BT}\right)\right], \quad (40)$$

with the prefactor[39,40]

$$N_n(T) = \frac{1}{4}\left(\frac{2m_0m_n^*k_BT}{\pi\hbar^2}\right)^{3/2}. \quad (41)$$

In the case of the satellite CB valleys at the L- and X-points (see below), m_n^* is the density-of-states mass for a single valley.

Similarly, the light hole density is given by[42,43]

$$p_{lh}(T) = N_{lh}(T)\left[F_{\frac{1}{2}}\left(-\frac{\mu}{k_BT}\right) + \right. \\ \left. + \frac{15}{4}\alpha_{lh}k_BT F_{\frac{3}{2}}\left(-\frac{\mu}{k_BT}\right)\right]. \quad (42)$$

For the heavy hole band, we do not consider the nonparabolicity[20] and set α_{hh}=0. We fix the heavy hole mass at m_{hh}=0.43, independent of temperature. This mass is determined by the separation E_0' between the p-bonding VB and the p-antibonding CB at the Γ-point, which has a weak relative temperature dependence.[44] This results in[42]

$$p_{hh}(T) = N_{hh}(T) F_{\frac{1}{2}}\left(-\frac{\mu}{k_BT}\right). \quad (43)$$

For completeness, we also add additional terms to consider the possibility of holes occupying the split-off hole band and electrons occupying the higher conduction band valleys at the L- and X-points:[42]

$$p_{so}(T) = N_{so}(T) F_{\frac{1}{2}}\left(\frac{-\Delta_0-\mu}{k_BT}\right), \quad (44)$$

$$n_L(T) = 4N_L(T) F_{\frac{1}{2}}\left(\frac{\mu-E_L^{\mathrm{exp}}}{k_BT}\right), \quad (45)$$

$$n_X(T) = 3N_X(T) F_{\frac{1}{2}}\left(\frac{\mu-E_X^{\mathrm{exp}}}{k_BT}\right). \quad (46)$$

The mass m_{so}^* calculated using Eq. (7) equals 0.15 at low temperatures, which is within the range of values given in the literature.[1,6] The dominant contribution to m_{so}^* comes from the spin-orbit splitting Δ_0 and therefore the smaller gap E_0^u at 800 K causes only a slight reduction of m_{so}^* to 0.14. For the positions of the satellite valleys at the L- and X-points, we use $\Delta_{\Gamma L}$=0.51 eV and $\Delta_{\Gamma X}$=0.83 eV, both with a density-of-states mass for a single valley of m_L=m_X=0.25, independent of temperature. We assume that these valleys shift rigidly with

temperature at the same rate as $E_0^{\rm exp}$. There are four L-valleys and three X-valleys in zinc blende semiconductors. (Diamond-type semiconductors have six Δ-valleys due to the double degeneracy at the X-point caused by the nonsymmorphic diamond space group.) Since little is known about the satellite CB valleys in InSb, these numbers are not much more than an educated guess. The carrier densities in the split-off hole band and in the X-valleys are negligible, but 10% of electrons occupy the L-valleys at 800 K. This was not considered in the analysis of the Hall experiments by Oszwaldowski and Zimpel,[23] as far as we know. Percentages of the electron and hole populations in the various bands are shown in the supplementary material.[13]

IV. RESULTS AND DISCUSSION

We can find the chemical potential of an intrinsic semiconductor from the charge neutrality condition[45]

$$n_\Gamma(T) + n_L(T) + n_X(T) = p_{lh}(T) + p_{hh}(T) + p_{so}(T) \tag{47}$$

at a given temperature T, for example using polylogarithm functions[12,38] in MATLAB.[11] As an example, we show the electron and hole density of InSb at 300 K as a function of the chemical potential in Fig. 7. At this temperature, the experimental "thermal" band gap $E_0^{\rm exp}$=0.187 eV and the unrenormalized mass band gap is E_0^u=0.221 eV. The room-temperature effective masses are m_e^*=0.0138 and m_{lh}^*=0.0142, calculated using Eqs. (10,11) in the large SO splitting approximation. For the holes, the light hole density is only a very small contribution, because the heavy hole is about 30 times heavier than the light hole. Therefore, the nonparabolicity correction does not matter much for the hole bands. The nonparabolicity correction for the electron concentration is sizeable, which can be seen from the difference between the red dotted and solid lines. The effective electron mass becomes larger at higher energies as shown by Eq. (36) and therefore the electron density is larger than in the parabolic case, because the prefactor Eq. (41) is proportional to $m_e^{1.5}$.

For a given temperature T, we plot n and p as a function of μ. The intrinsic chemical potential is found at the location where the two lines cross,[50] thus satisfying the charge neutrality condition (47). In the parabolic case, the electron and hole densities versus chemical potential cross at E_F=162 meV. In the nonparabolic case, the electron and hole densities cross at a lower chemical potential of E_F=157 meV because of the larger electron density. At 300 K, the Fermi level is just below the bottom of the CB, because $E_0^{\rm exp}$=0.187 eV as mentioned earlier. The intrinsic carrier concentration of InSb at 300 K is 13.6×10^{15} cm^{-3} for parabolic bands and 16.4×10^{15} cm^{-3} in the nonparabolic case.

This method is used to find the chemical potential at each temperature, as shown in Fig. 8. (Compare Fig. 1 of Masut.[7]) At low temperatures, the chemical potential is approximately equal to half the band gap[39] and therefore the argument of the Fermi integral is very small. For this case, we can apply the nondegenerate limit[46–48]

$$F_{\frac{1}{2}}(\eta) \approx \exp(\eta) \quad \text{for} \quad \eta \ll -1, \tag{48}$$

essentially using classical Maxwell-Boltzman statistics to describe the electron and hole populations. This approx-

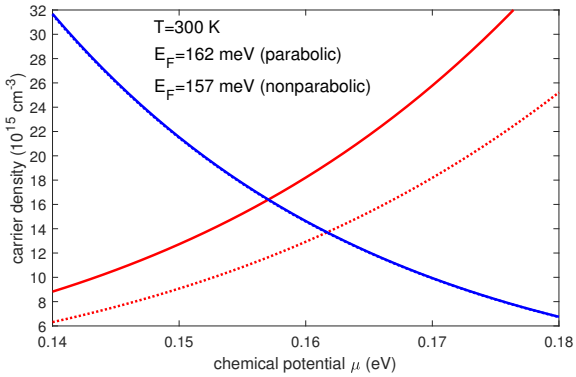

FIG. 7. Electron (red) and hole density (blue) of InSb as a function of chemical potential at 300 K in the parabolic approximation (dotted) and with the lowest nonparabolic corrections (solid). The thermal and mass band gaps were taken from Fig. 4 and the electron and light hole masses were calculated in the large SO approximation.

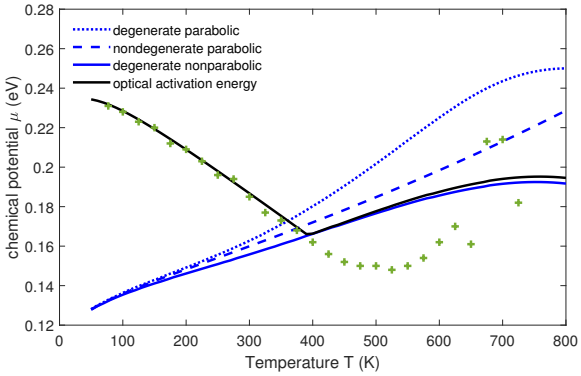

FIG. 8. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large SO approximation. The full temperature dependence of the direct gap according to Eq. (28) was included in the Fermi-Dirac integral, but the effective masses were calculated taking into account only the thermal expansion contribution to the band gap given by Eq. (24), not the electron-phonon renormalization. The optical activation energy from Eq. (52) and the experimental direct band gap[38] from a fit to the temperature-dependent infrared dielectric function with a Johs-Herzinger parametric oscillator model (symbols) are also shown.

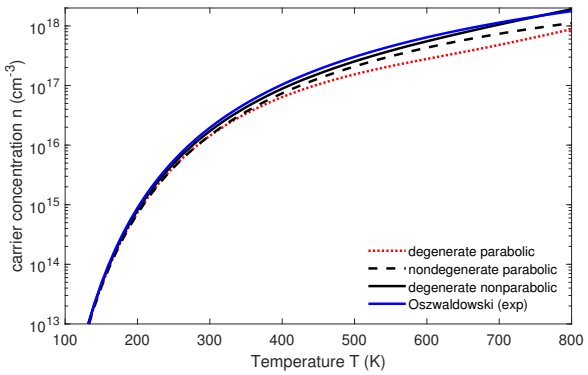

FIG. 9. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest nonparabolic correction in the large SO approximation. The temperature dependence of the direct gap according to Eq. (28) was included in the Fermi-Dirac distribution function, but the effective masses were calculated taking into account only the thermal expansion contribution to the band gap from Eq. (24), not the electron-phonon renormalization. The blue line shows a fit to carrier concentrations determined from Hall measurements by Oszwaldowski and Zimpel.[23]

imation leads to the well-known expressions[39,40]

$$\mu \approx \frac{E_0}{2} + \frac{3}{4} k_B T \ln\left(\frac{m^*_{hh}}{m^*_e}\right) \quad \text{and} \tag{49}$$

$$n \approx 2\left(\frac{m_0 k_B T}{2\pi\hbar^2}\right)^{\frac{3}{2}} (m^*_e m^*_{hh})^{\frac{3}{4}} \exp\left(-\frac{E_0}{2k_B T}\right). \tag{50}$$

As shown in Fig. 8, these nondegenerate expressions can be used up to 300 K for InSb, but deviations become noticeable at higher temperatures. The chemical potential increases nearly linearly with temperature below 300 K as implied by Eq. (49). The small deviation from linearity is caused by the temperature dependence of the effective electron mass. Above 300 K, we must evaluate the Fermi-Dirac integral exactly using polylogarithm functions.[12] The fully degenerate limit (where the argument of the Fermi-Dirac integral is very large) is never reached for intrinsic InSb.

In general, degenerate Fermi-Dirac statistics leads to a higher chemical potential than non-degenerate (classical) Maxwell-Boltzmann statistics, as shown by the comparison for parabolic bands. Including the nonparabolicity significantly reduces the chemical potential, as we have already seen in Fig. 7.

The intrinsic carrier concentration as a function of temperature is shown in Fig. 9. For parabolic bands, considering degenerate carrier statistics reduces the carrier density.[50] If nonparabolic bands are considered, then the effective electron mass becomes larger, which increases the carrier concentration according to Eq. (50). Oszwaldowski and Zimpel[23] obtained the temperature dependence of the intrinsic carrier concentration of InSb from 200 to 800 K with Hall measurements. Assuming a Hall factor of unity, they found an intrinsic carrier concentration near 1.8×10^{18} cm^{-3} at 800 K. They fitted their results with the expression[23]

$$n = 2.9\times10^{11}\,(2400-T)^{0.75}\left(1+2.7\times10^{-4}T\right)T^{1.5} \times \exp\left(-\frac{0.129-1.5\times10^{-4}T}{k_B T}\right), \tag{51}$$

where n is in units of cm^{-3}, T in K, and $k_B T$ in eV. This Hall concentration is also shown in Fig. 9. Our calculation finds a carrier concentration of 1.9×10^{18} cm^{-3} at 800 K, but this agreement is better than it should be. Our use of the large SO splitting approximation overestimates the effective electron mass by 12% at 0 K and by 10% at 800 K. According to Eq. (50), our model should also overestimate the carrier concentration. Another uncertainty in our model is the temperature dependence of the heavy hole mass, which has been discussed in the literature to a good extent.[23]

We also compare the results for the chemical potential in Fig. 8 with optical measurements of the band gap.[38] The optical activation energy (i.e., the band gap observed in an optical absorption or ellipsometry experiment) is increased through the Burstein-Moss shift and given by[49]

$$E_A = \max\left[E_0, E_0 + \left(1+\frac{m^*_e}{m^*_{hh}}\right)(\mu - E_0)\right]. \tag{52}$$

The optical activation energy is equal to E_0 if the Fermi level is below the conduction band minimum, but increases as the Fermi level moves into the conduction band above 400 K. The ratio of the masses takes into account that direct optical interband transitions are not possible at k=0, if the Fermi level is larger than the band gap. This optical activation energy is also shown in Fig. 8. It qualitatively describes the upward trend of the ellipsometry data of Rivero Arias *et al.*[38] shown by symbols at higher temperatures.

V. SUMMARY

We have shown how a simple 8×8 $\vec{k}\cdot\vec{p}$ model due to Kane[2] within the large spin-orbit splitting approximation can be used to describe the nonparabolicity of the light hole and conduction bands in cubic diamond and zinc blende semiconductors at the Γ-point. This model treats the interaction of the p-bonding valence bands with the s-antibonding conduction band with a single parameter E_P, which is related to the momentum matrix element. The band gap E_0 and the spin-orbit splitting Δ_0 are the other two parameters of the model. As an application, we have derived analytical expressions for the effective electron and light hole masses, the chemical potential, and the carrier concentration of intrinsic InSb as a function of temperature. The results are in excellent agreement with Hall measurements of the carrier concentration[23] if the

unrenormalized band gap E_0^u is used to calculate the effective masses. E_0^u includes the contribution of thermal expansion to the temperature dependence of the band gap, but not its renormalization due to electron-phonon interactions. The replacement of the experimental band gap E_0 by the unrenormalized gap E_0^u requires a small adjustment of the momentum matrix element.

ACKNOWLEDGMENTS

This research was supported in part by the Air Force Research Laboratory Sensors Directorate, through the Air Force Office of Scientific Research Summer Faculty Fellowship Program®, Contract Nos. FA8750-15-3-6003, FA9550-15-0001 and FA9550-20-F-0005. This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-20-1-0135. This material is based upon work supported by the National Science Foundation under Award No. (DMR-2235447).

[1]P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 2010).
[2]E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957).
[3]M. Cardona, N. E. Christensen, and G. Fasol, Phys. Rev. Lett. **56**, 2831 (1986).
[4]M. Fox, *Optical Properties of Solids* (Oxford University Press, Oxford, 2010).
[5]E. O. Kane, in *Semiconductors and Semimetals*, vol. 1, edited by R. K. Willardson and A. C. Beer, p. 75 (Academic, New York, 1966).
[6]P. Lawaetz, Phys. Rev. B **4**, 3460 (1971).
[7]R. A. Masut, Eur. J. Phys. **43**, 015501 (2022).
[8]W. Zawadzki and J. Kolodziejczak, Phys. Status Solidi **6**, 409 (1964)
[9]W. Zawadzki, R. Kowalczyk, and J. Kolodziejczak, Phys. Status Solidi **10**, 513 (1965).
[10]W. Zawadzki, Adv. Phys. **23**, 435 (1974).
[11]`http://www.mathworks.com`.
[12]M.D. Ulrich, W.F. Seng, and P.A. Barnes, J. Comp. Electron. **1**, 431 (2002).
[13]See supplementary material at *URL to be inserted by AIP Publishing* for detailed derivations and MATLAB scripts to perform the calculations presented here.
[14]E. O. Kane, J. Phys. Chem. Solids **1**, 82 (1956).
[15]W. Shockley, Phys. Rev. **78**, 173 (1950).
[16]G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. **98**, 368 (1955).
[17]M. Cardona and F. H. Pollak, Phys. Rev. **142**, 530 (1966).
[18]R. W Cunningham and J. B. Gruber, J. Appl. Phys. **41**, 1804 (1970).
[19]E. M. Conwell and M. O. Vassell, Phys. Rev. **166**, 797 (1968). Note that this reference addresses the small SO limit $E_0 \gg \Delta_0$.
[20]F. J. Bartoli, J. R. Meyer, C. A. Hoffmann, and R. E. Allen, Phys. Rev. B **27**, 2248 (1983).
[21]C. Jacoboni and L. Reggiani, Rev. Mod. Phys. **55**, 645 (1983).
[22]C. Emminger, N. S. Samarasingha, M. Rivero Arias, F. Abadizaman, J. Menendez, and S. Zollner, J. Appl. Phys. **131**, 165701 (2022).
[23]M. Oszwaldowski and M. Zimpel, J. Phys. Chem. Solids **49**, 1179 (1988).
[24]E. Burstein, Phys. Rev. **93**, 632 (1954).
[25]T. S. Moss, Proc. Phys. Soc. B **67**, 775 (1954).
[26]P. Vashista and R. K. Kalia, Phys. Rev. B **25**, 6492 (1982)
[27]R. Zimmerman, Phys. Status Solidi B **146**, 371 (1988)
[28]M. Cardona and S. Gopalan, in *Progress in Electron Properties of Solids*, edited by E. Doni, R. Girlanda, G. P. Parravicini, and A. Quattropani (Springer, Dordrecht, 1989), p. 51.
[29]S. Zollner, S. Gopalan, and M. Cardona, Solid State Commun. **77**, 485 (1991).
[30]S. Zollner, M. Garriga, J. Humlíček, S. Gopalan, and M. Cardona, Phys. Rev. B **43**, 4349 (1991).
[31]S. I. Novikova, in *Semiconductors and Semimetals*, vol. 2, edited by R. K. Willardson and A. C. Beer, p. 33 (Academic, New York, 1966).
[32]X. Cai and J. Wei, J. Appl. Phys. **114**, 083507 (2013).
[33]D. F. Gibbons, Phys. Rev. **112**, 136 (1958).
[34]P. W. Sparks and C. A. Swenson, Phys. Rev. **163**, 779 (1967).
[35]P. E. Van Camp, V. E. Van Doren, and J. T. Devreese, Phys. Rev. B **41**, 1598 (1990).
[36]I. M. Booth, M. H. Hawton, and W. J. Keeler, Phys. Rev. B **25**, 7713 (1982).
[37]R. Roucka, Y.-Y. Fang, J. Kouvetakis, A. V. G. Chizmeshya, and J. Menéndez, Phys. Rev. B **81**, 245214 (2010).
[38]M. Rivero Arias, C. A. Armenta, C. Emminger, C. M. Zamarripa, N. S. Samarasingha, J. Love, S. Yadav, and S. Zollner, J. Vac. Sci. Technol. B **41**, 022203 (2023).
[39]N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders, Philadelphia, 1976).
[40]S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981).
[41]H. Ehrenreich, Phys. Rev. **120**, 1951 (1960).
[42]J. Menéndez, C. D. Poweleit, and S. E. Tilton, Phys. Rev. B **101**, 195204 (2020).
[43]C. A. Hoffmann, J. R. Meyer, R. J. Wagner, F. J. Bartoli, M. A. Engelhardt, and H. Höchst, Phys. Rev. B **40**, 11693 (1989).
[44]S. Logothetidis, L. Viña, and M. Cardona, Phys. Rev. B **31**, 947 (1985).
[45]J. S. Blakemore, *Semiconductor Statistics* (Pergamon, Oxford, 1962).
[46]J. McDougall and E. C. Stoner, Philos. Trans. R. Soc. London **237**, 67 (1938).
[47]J. S. Blakemore, Solid-State Electron. **25**, 1067 (1982).
[48]R. Kim, X. Wang, and M. Lundstrom, *Notes on Fermi-Dirac Integrals (4th Edition)*, `https://nanohub.org/resources/5475` (2008).
[49]P. K. Chakraborty, G. C. Datta, K. P. Ghatak, Physica B **339**, 198 (2003).
[50]W. Shockley, *Electrons and Holes in Semiconductors* (Van Nostrand, Princeton, 1950).

Supplementary Material for "Conduction band nonparabolicity, chemical potential, and carrier concentration of intrinsic InSb as a function of temperature"

(Dated: 29 December 2023)

Stefan Zollner,[1] Carlos A. Armenta,[1] Sonam Yadav,[1] and José Menéndez[2]

[1] *Department of Physics, New Mexico State University, MSC 3D, P.O. Box 30001, Las Cruces, NM 88003, USA*
[2] *Department of Physics, Arizona State University, Tempe, AZ 85287-1504, USA*

S1. QUANTUM MECHANICAL PERTURBATION THEORY

A. Non-degenerate perturbation theory

Let us assume [following the Wikipedia article *Perturbation theory (quantum mechanics)*] that a quantum mechanical system is described, for the most part, by a Hamiltonian H_0 which has a finite number of eigenstates $|n0\rangle$ with eigenvalues E_{n0} given by

$$H_0 \,|n0\rangle = E_{n0} \,|n0\rangle \,; \quad n = 1, \ldots, N. \tag{S1}$$

In addition, there is a small perturbation $H_{\vec{k}}$ that depends on some vector parameter $\vec{k}$, which can be made arbitrarily small. The total Hamiltonian is therefore $H = H_0 + H_{\vec{k}}$ with eigenstates $\left|n\vec{k}\right\rangle$ and eigenvalues $E_{n\vec{k}}$ given by

$$H \left|n\vec{k}\right\rangle = E_{n\vec{k}} \left|n\vec{k}\right\rangle ; \quad n = 1, \ldots, N. \tag{S2}$$

To calculate the perturbed eigenvalues and eigenstates of H for small values of $\left|\vec{k}\right|$, we must distinguish between two separate cases, depending on whether the eigenstate $|n0\rangle$ is non-degenerate or degenerate.

In the **non-degenerate** case, the perturbed eigenvalues and eigenstates are given to **first order** in the perturbation by

$$E_{n\vec{k}} \;=\; E_{n0} + \langle n0| \, H_{\vec{k}} \, |n0\rangle \,, \tag{S3}$$

$$\left|n\vec{k}\right\rangle \;=\; |n0\rangle + \sum_{i\neq n} |i0\rangle \, \frac{\langle i0| \, H_{\vec{k}} \, |n0\rangle}{E_{n0} - E_{i0}}. \tag{S4}$$

We sometimes also need to know the eigenvalues to second order in $\left|\vec{k}\right|$, which are given in **second order** perturbation theory by

$$E_{n\vec{k}} = E_{n0} + \langle n0| \, H_{\vec{k}} \, |n0\rangle + \sum_{i\neq n} \frac{\left|\langle i0| \, H_{\vec{k}} \, |n0\rangle\right|^2}{E_{n0} - E_{i0}}. \tag{S5}$$

These equations can be written in more compact form, if we introduce the notation

$$H_{ni} = \langle n0| \, H_{\vec{k}} \, |i0\rangle \quad \text{and} \quad E_{ni} = E_{n0} - E_{i0}. \tag{S6}$$

The first order perturbation theory results in

$$E_{n\vec{k}} \;=\; E_{n0} + H_{nn}, \tag{S7}$$

$$\left|n\vec{k}\right\rangle \;=\; |n0\rangle + \sum_{i\neq n} |i0\rangle \, \frac{H_{in}}{E_{ni}} \tag{S8}$$

and second order perturbation theory in

$$E_{n\vec{k}} = E_{n0} + H_{nn} + \sum_{i\neq n} \frac{|H_{ni}|^2}{E_{ni}}. \tag{S9}$$

B. Degenerate perturbation theory

If the energy level E_{n0} is **degenerate**, then Eq. (S3) is ambiguous, because H_{nn} depends on the choice of the basis and therefore may not be uniquely defined. Also, the energy denominator E_{ni} will vanish if $|i0\rangle$ is in the degenerate eigenspace belonging to eigenvalue E_{n0} and therefore the second-order energy correction (S5) will become singular.

We address this by assuming that g is the degeneracy of the E_{n0} eigenspace of the unperturbed Hamiltonian. If we choose arbitrary basis functions $|i0\rangle$ for this eigenspace, then the wave function $|n0\rangle$ can be written as a linear combination

$$|n0\rangle = \sum_{i=1}^{g} c_i \,|i0\rangle \tag{S10}$$

with some coefficients c_i. The perturbed energies to first order perturbation theory then become

$$E_{n\vec{k}} = E_{n0} + E_i, \tag{S11}$$

where the E_i are the eigenvalues of the perturbation matrix $H_{ij} = \langle i0| \, H_{\vec{k}} \, |j0\rangle$ within the degenerate eigenspace of E_{n0}. The perturbation will usually lift some, but not necessarily all of the degeneracies of this eigenspace.

If the perturbation is sizeable and the corrected energies encroach on nearby energy levels of the unperturbed Hamiltonian, it may be necessary to also include the nearby levels in the calculation. In first-order perturbation theory, the perturbed energies are then given by the eigenvalues of the matrix

$$\langle i0| \, H_0 + H_{\vec{k}} \, |j0\rangle \,, \tag{S12}$$

where as many eigenstates of H_0 as needed can be included in the matrix.

S2. $\vec{k}\cdot\vec{p}$-THEORY

Electron states in solids are described by three quantum numbers, the wave vector $\vec{k}$, the band index n with energy $E_{n\vec{k}}$, and the spin s (which we ignore for the moment). The one-electron Hamiltonian H is the sum of the kinetic energy derived from the momentum operator $\vec{p}=-i\hbar\vec{\nabla}$ and the crystal potential $V(\vec{r})$. If $\phi_{n\vec{k}}$ is the wave function, then the Schrödinger equation is

$$H\phi_{n\vec{k}}=\left[\frac{\vec{p}^2}{2m_0}+V(\vec{r})\right]\phi_{n\vec{k}}=E_{n\vec{k}}\phi_{n\vec{k}}, \tag{S13}$$

where m_0 is the free electron mass. Bloch's theorem allows us to write the wave function as the product of a plane wave and a function $u_{n\vec{k}}$, which is periodic in the crystal lattice:

$$\phi_{n\vec{k}}=e^{i\vec{k}\cdot\vec{r}}u_{n\vec{k}}. \tag{S14}$$

If we insert Eq. (S14) into the Schrödinger equation (S13) and use the product rule for the Laplace operator $\Delta fg=f\Delta g+2\vec{\nabla}f\cdot\vec{\nabla}g+g\Delta f$, we obtain a new equation for the periodic part of the wave function, see Yu and Cardona (2.35):

$$\left[\frac{\vec{p}^2}{2m_0}+\frac{\hbar}{m_0}\vec{k}\cdot\vec{p}+\frac{\hbar^2k^2}{2m_0}+V(\vec{r})\right]u_{n\vec{k}}=E_{n\vec{k}}u_{n\vec{k}}. \tag{S15}$$

For a vanishing wave vector $\vec{k}$=0, the kinetic energy and potential terms have solutions that are assumed to be known, for example from optical measurements of the gaps at critical points, see Yu & Cardona (2.36),

$$\left[\frac{\vec{p}^2}{2m_0}+V(\vec{r})\right]u_{n0}=E_{n0}u_{n0}. \tag{S16}$$

The $\vec{k}$-dependent terms in Eq. (S15) are treated in perturbation theory up to second order in $\vec{k}$.

S3. KANE'S SUBSTITUTION

The physical significance of the four terms in square brackets in Eq. (S15) is as follows: The first and fourth terms are the kinetic and potential energy of the electron at the Γ-point ($\vec{k}$=0) belonging to the unperturbed Hamiltonian. The second term is the $\vec{k}\cdot\vec{p}$-term, which will be treated in perturbation theory. The third term, which is quadratic in k, is called the free electron term, because it describes the kinetic energy of a free electron with mass m_0 as a function of crystal momentum $\hbar\vec{k}$.

Kane (1956, 1957) notes that the formalism can be simplified by introducing a **modified energy parameter**

$$\tilde{E}_{n\vec{k}}=E_{n\vec{k}}-\frac{\hbar^2k^2}{2m_0}, \tag{S17}$$

where the free electron kinetic energy has been subtracted. We use a tilde (instead of a prime) to denote the modified energies, because the prime (as in E_0') has a different meaning in the critical-point notation proposed by Cardona.

This simplifies Eq. (S15) to

$$\left[\frac{\vec{p}^2}{2m_0}+\frac{\hbar}{m_0}\vec{k}\cdot\vec{p}+V(\vec{r})\right]u_{n\vec{k}}=\tilde{E}_{n\vec{k}}u_{n\vec{k}}. \tag{S18}$$

The $\vec{k}\cdot\vec{p}$-term is then treated as a perturbation.

For a non-degenerate band, the first-order correction is

$$\tilde{E}_{n\vec{k}}=E_{n\vec{k}}-\frac{\hbar^2k^2}{2m_0}=E_{n0}+\frac{\hbar}{m_0}\langle n0|\,\vec{k}\cdot\vec{p}\,|n0\rangle. \tag{S19}$$

The term with the dot product is linear in $\vec{k}$. This term must vanish in crystals with inversion symmetry. It also vanishes in crystals without inversion symmetry due to the Kramers degeneracy (time-reversal symmetry), if spin effects are neglected. Note: There are small $\vec{k}$-linear terms in the band structure of zinc blende semiconductors, but they arise from relativistic corrections, not from the last term in Eq. (S19).

The perturbed wave function for non-degenerate bands is derived from Eq. (S4) and found to be, compare Yu & Cardona (2.37),

$$u_{n\vec{k}}=u_{n0}+\frac{\hbar}{m_0}\sum_{i\neq n}|i0\rangle\frac{\langle i0|\,\vec{k}\cdot\vec{p}\,|n0\rangle}{E_{n0}-E_{i0}}. \tag{S20}$$

The perturbed wave functions in Eq. (S20) are needed to calculate the correction to the energy in second order perturbation theory with Eq. (S5).

Since the first-order energy correction vanishes, we proceed with the correction in second-order perturbation theory and find

$$\tilde{E}_{n\vec{k}}=E_{n0}+\frac{\hbar^2}{m_0^2}\sum_{i\neq n}\frac{\left|\langle n0|\,\vec{k}\cdot\vec{p}\,|i0\rangle\right|^2}{E_{n0}-E_{i0}} \tag{S21}$$

for non-denerate levels E_{n0}.

Similarly, we need to diagonalize the matrix

$$\sum_i\left(E_{n0}\delta_{ni}+\frac{\hbar}{m_0}\langle n0|\,\vec{k}\cdot\vec{p}\,|i0\rangle\right)c_{ni}=\tilde{E}_{n\vec{k}}c_{nn} \tag{S22}$$

in the degenerate case. Note that the energies on the left hand side in the diagonal of the Hamiltonian are the unperturbed eigenvalues E_{n0} while the energies on the right hand side are Kane's modified energies $\tilde{E}_{n\vec{k}}$. This follows from Eq. (S18). The perturbed wave functions at finite $\vec{k}$ for the degenerate case are given by, see Kane 1966, Eq. (4),

$$u_{n\vec{k}}=\sum_i c_{in}u_{i0}, \tag{S23}$$

expressed in terms of the unperturbed wave functions at $\vec{k}$=0.

For specificity and for comparison with the literature, we rewrite equations (S21-S22) by explicitly adding the free-electron term: For a **non-degenerate** band with energy E_{n0}, the correction to the energy up to second order in $\vec{k}$ (and ignoring the linear terms) is, compare Yu & Cardona (2.38),

$$E_{n\vec{k}} = E_{n0} + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{i \neq n} \frac{\left| \langle n0| \, \vec{k} \cdot \vec{p} \, |i0\rangle \right|^2}{E_{n0} - E_{i0}}. \quad \text{(S24)}$$

For **degenerate** bands, following the eigenvalue problem posed in Eq. (S12), the energies for finite $\vec{k}$ are found by diagonalizing the matrix, see Kane (1966), Eq. (8),

$$\sum_i \left[\left(E_{n0} + \frac{\hbar^2 k^2}{2m_0} \right) \delta_{ni} + \frac{\hbar}{m_0} \langle n0| \, \vec{k} \cdot \vec{p} \, |i0\rangle \right] c_{ni} = E_{n\vec{k}} c_{nn}. \quad \text{(S25)}$$

The eigenvalues of this matrix are the band energies $E_{n\vec{k}}$ near $\vec{k}$=0 from $\vec{k} \cdot \vec{p}$-theory, including the free-electron energy.

Kane's substitution (S17) only changes the energies, but it does not affect the wave functions. Therefore, Eqs. (S20) and (S23) are valid either way.

It is common to include 3, 4, or 7 bands in the $\vec{k} \cdot p$ matrix. A 3×3-matrix is needed to describe the top valence bands of Si and Ge (see, for example, Kane 1956). With a 4×4-matrix, one can also include the the lowest conduction band (see Kane 1957). A 7×7-matrix includes the interaction between the p-bonding valence bands and the p-antibonding conduction bands. This interaction is needed to describe the warping of the valence bands (Dresselhaus, Kip, Kittel). Adding the s-bonding valence band as well completes an 8×8-matrix. Including higher-energy d-orbitals in the conduction band further increases the size of the matrix (Cardona & Pollak 1966). The rank of this matrix is doubled if the spin degeneracy is included explicitly. For example, a commonly used 14×14-matrix includes the p-type valence band and the s-type and p-type conduction bands with spin.

S4. APPLICATION TO THE NON-DEGENERATE CONDUCTION BAND IN INSB

We can apply Eq. (S24) to the lowest conduction band of InSb at the Γ-point. The (unperturbed) energy of this band at $\vec{k} = 0$ is called E_0. The strongest contribution to Eq. (S24) comes from the top valence bands, because they have the smallest energy denominator. The top valence band is threefold degenerate with basis functions X, Y, and Z, see Yu & Cardona (2.50). The only non-vanishing matrix elements in Eq. (S24) are of the form, see Yu & Cardona (2.42),

$$\langle s| \, p_x \, |X\rangle = -iP. \quad \text{(S26)}$$

The mixed momentum matrix elements $\langle s| \, p_x \, |Y\rangle$ etc vanish. Without loss of generality, we can assume that the $\vec{k}$-vector points along the z-direction. For other directions, one can rotate the coordinate system using Eqs. (6-9) in Kane 1957. Ignoring the contributions from all other "remote" bands, the conduction band energy near the minimum (for small $\vec{k}$) is found to be

$$\begin{aligned} E_{\mathrm{CB}}\left(\vec{k}\right) &= E_0 + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2 k^2}{m_0^2} \frac{P^2}{E_0} = \\ &= E_0 + \frac{\hbar^2 k^2}{2m_0} \left(1 + \frac{2}{m_0} \frac{P^2}{E_0} \right), \end{aligned} \quad \text{(S27)}$$

if we neglect the spin-orbit splitting Δ_0. By selecting the $\vec{k}$-vector along one of the axes, the sum in Eq. (S24) only has a single term.

If we wish to include the spin-orbit splitting Δ_0 of the valence band maximum, then we need to be careful and use the six basis functions for the valence band given by Eqs. (2.48) and (2.49) in Yu & Cardona instead of the X, Y, Z basis set. See also Eq. (32) in Kane (1966). Still selecting $\vec{k}$ along the z-axis, the sum now has two terms, one for the j=$\frac{3}{2}$ state and one for the j=$\frac{1}{2}$ state. Spin doubles the number of states, but matrix elements between different spin states vanish. The result of Eq. (S24) is

$$E_{\mathrm{CB}}\left(\vec{k}\right) = E_0 + \frac{\hbar^2 k^2}{2m_0} \left[1 + \frac{2P^2}{3m_0} \left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0} \right) \right], \quad \text{(S28)}$$

compare Eq. (12) in Kane (1957) or Eq. (5) in Cardona (1963) or Eq. (11) in Emminger (2022). The factor $\frac{1}{3}$ in this equation is derived from the normalization of the basis functions.

We introduce the energy

$$E_P = \frac{2P^2}{m_0}, \quad \text{(S29)}$$

which equals about 20 to 25 eV for many common semiconductors. The conduction band energy then becomes

$$E_{\mathrm{CB}}\left(\vec{k}\right) = E_0 + \frac{\hbar^2 k^2}{2m_0} \left(1 + \frac{2}{3} \frac{E_P}{E_0} + \frac{1}{3} \frac{E_P}{E_0 + \Delta_0} \right), \quad \text{(S30)}$$

We recognize the term in parentheses as the inverse effective mass of the lowest conduction band

$$\frac{1}{m_e^*} = 1 + \frac{2}{3} \frac{E_P}{E_0} + \frac{1}{3} \frac{E_P}{E_0 + \Delta_0} = 1 + \frac{E_P}{3} \left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0} \right), \quad \text{(S31)}$$

compare Emminger (2022), Eq. (11). We stress that this result was obtained using second-order perturbation theory for non-degenerate energy states, keeping only terms up to second order in $\vec{k}$. The conduction band (S30) turns out to be parabolic, because remote bands and higher order terms in $\vec{k}$ were omitted. Using the values E_0=0.237 eV, Δ_0=0.81 eV, and E_P=23.1 eV at low temperature (Lawaetz 1971), we find m_e=0.0136 for InSb, which compares well to the experimental value of 0.0137. See also Table 2.22 in Yu & Cardona.

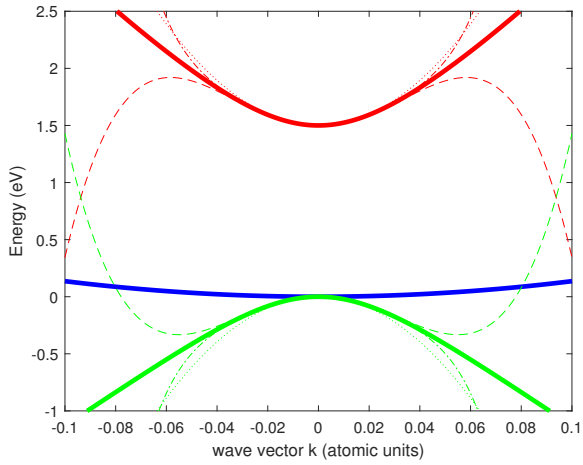

FIG. S1. Band structure of GaAs near the Brillouin zone center calculated from a 4×4 $\vec{k}\cdot\vec{p}$-model within the approximation for small spin-orbit splitting. The electron, light hole, and heavy hole bands are shown in red, green, and blue, respectively. Thick lines show the dispersion calculated from Eqs. (S35) and (S36) and dotted lines the parabolic approximation. Dashed and dot-dashed lines include the corrections proportional to k^4 and k^6 given by Eq. (S45). The wave vector is shown in atomic units (inverse Bohr radii).

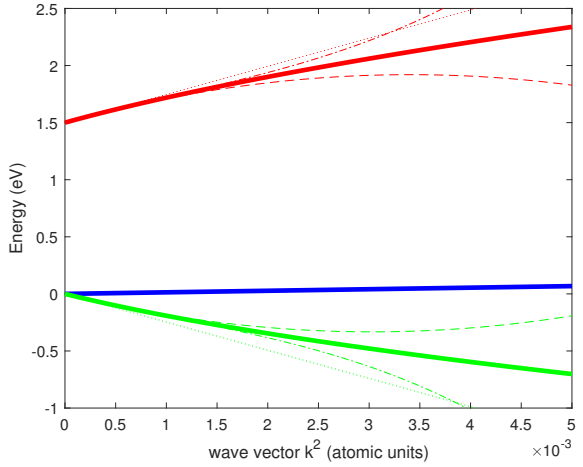

FIG. S2. As Fig. S1, but as a function of the square of the wave vector in atomic units, to better show the deviations from linearity.

S5. BAND ENERGIES FROM 4×4 $\vec{k}\cdot\vec{p}$-MODEL WITHOUT SPIN-ORBIT COUPLING

To understand the origin of the nonparabolicity of the valence and conduction bands, we start our analysis by evaluating the matrix given in Eq. (S22) with a 4×4 $\vec{k}\cdot\vec{p}$-model, which includes the three degenerate top valence bands and the lowest conduction band. We first ignore spin-orbit splitting by setting Δ_0=0. (This will be changed later for InSb, where the spin-orbit splitting is large.) We use the basis set $|S\rangle$, $|X\rangle$, $|Y\rangle$, and $|Z\rangle$. The only non-vanishing momentum matrix elements are those given by Eq. (S26). The energy reference level E_v=0 is taken to be at the top of the valence band. Without loss of generality, we select the $\vec{k}$-vector along the x-direction. For other directions, one can rotate the coordinate system using Eqs. (6-9) in Kane 1957.

The $\vec{k}\cdot\vec{p}$-matrix for the modified energies (S17,S22) is

$$H_0+\tilde{H}_{\vec{k}}=\begin{pmatrix} E_0 & -\frac{\hbar k}{m_0}iP & 0 & 0\\ \frac{\hbar k}{m_0}iP & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}, \qquad \text{(S32)}$$

where the top row and the left column represent the $|S\rangle$-state and the bottom row and right column the $|Z\rangle$-state. The tilde over $H_{\vec{k}}$ indicates that the free-electron kinetic energy has been subtracted as indicated in Eq. (S17). The matrix (S32) is the same as Eq. (4) in Kane (1957), if we set Δ_0=0 and consider that Kane used $|iS\rangle$ as one of the basis functions, which makes the off-diagonal elements real. Compare also Eq. (20) in Kane (1966), which includes many higher-order matrix elements that we have set to zero. But note that Kane (1966) does not introduce the modified energies (S17) with a tilde and instead adds the k^2 term explicitly to the diagonal elements.

This matrix has two eigenvalues $\tilde{E}_{1,2}$=0, compare Eq. (9) in Kane (1957). The other two eigenvalues can be found by solving the characteristic equation

$$\begin{aligned}\left(E_0-\tilde{E}\right)\left(0-\tilde{E}\right)-\frac{\hbar^2k^2P^2}{m_0^2} &= \\ =\tilde{E}\left(\tilde{E}-E_0\right)-\frac{\hbar^2k^2}{2m_0}E_P &= 0, \qquad \text{(S33)}\end{aligned}$$

which is equivalent to Kane's (1957) Eq. (10) and Kane's (1966) Eq. (38). The remaining two solutions for the modified energies are

$$\begin{aligned}\tilde{E}_{3,4} &= \frac{E_0}{2}\left[1\pm\sqrt{1+\left(\frac{2\hbar kP}{m_0E_0}\right)^2}\right]\\ &= \frac{E_0}{2}\left(1\pm\sqrt{1+4\frac{\hbar^2k^2}{2m_0}\frac{E_P}{E_0^2}}\right), \qquad \text{(S34)}\end{aligned}$$

which are also given by Kane (1966) in Eq. (43). By adding the k^2-term back in, see Eq. (S17), we obtain the four solutions for the energies in the conduction and valence band:

$$E_{1,2} = \frac{\hbar^2k^2}{2m_0} \qquad \text{(S35)}$$

$$\begin{aligned}E_{3,4} &= \frac{\hbar^2k^2}{2m_0}+\frac{E_0}{2}\left[1\pm\sqrt{1+\left(\frac{2\hbar kP}{m_0E_0}\right)^2}\right]=\\ &= \frac{\hbar^2k^2}{2m_0}+\frac{E_0}{2}\left(1\pm\sqrt{1+4\frac{\hbar^2k^2}{2m_0}\frac{E_P}{E_0^2}}\right). \qquad \text{(S36)}\end{aligned}$$

Equation (S36) can also be applied if the spin-orbit splitting is much smaller than the band gap, see for example Ehrenreich Phys. Rev. **120**, 1951 (1960), Eq. (1), or Conwell and Vassell, Phys. Rev. **166**, 797 (1968). Eq. (S36) can therefore be used for phosphides, the light arsenides AlAs and GaAs, and with limited accuracy perhaps even for Ge.

We can expand the square root for small values of k, which leads to the parabolic approximation

$$\begin{aligned} E_3 &= E_0 + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2 k^2 P^2}{m_0^2 E_0} = \\ &= E_0 + \frac{\hbar^2 k^2}{2m_0}\left(\frac{E_P}{E_0}+1\right), \qquad \text{(S37)} \\ E_4 &= \frac{\hbar^2 k^2}{2m_0} - \frac{\hbar^2 k^2 P^2}{m_0^2 E_0} = \\ &= \frac{\hbar^2 k^2}{2m_0}\left(1-\frac{E_P}{E_0}\right) = -\frac{\hbar^2 k^2}{2m_0}\left(\frac{E_P}{E_0}-1\right). \qquad \text{(S38)} \end{aligned}$$

We need to discuss these results: There are two degenerate bands $E_{1,2}$ with a positive mass of m_0, which curve upward. Compare Kane 1966, Sec. IV.3.b(2). These bands correspond to the heavy and split-off hole bands. The experimentally observed downward curvature of these bands is therefore caused by the interaction with remote bands at higher energies, which we have ignored in our 4×4-model, especially the matrix element Q, which describes the interaction between the p-bonding valence band and the p-antibonding conduction band through the DKK ABC parameters, see Emminger (2022), Eqs. (12-20). The dispersion of the band E_3 is the electron in the conduction band and the result is identical to Eq. (S30) for Δ_0=0. Finally, E_4 is the light hole band. It has a negative curvature, since the band gap E_0 is smaller than the momentum matrix element energy E_P. To stress the negative curvature, we have pulled a negative sign in front of the parentheses. This leads to a positive light hole mass.

The effective electron mass from Eq. (S37) is

$$\frac{1}{m_e^*} = 1 + \frac{E_P}{E_0} = \frac{E_P + E_0}{E_0} \approx \frac{E_P}{E_0} \qquad \text{(S39)}$$

or

$$m_e^* = \frac{E_0}{E_P + E_0} \approx \frac{E_0}{E_P}. \qquad \text{(S40)}$$

This is the same as Eq. (S31) for Δ_0=0. Similarly, the light hole mass obtained from Eq. (S38) is given by

$$\frac{1}{m_{lh}^*} = \frac{E_P}{E_0} - 1 = \frac{E_P - E_0}{E_0} \qquad \text{(S41)}$$

or

$$m_{lh}^* = \frac{E_0}{E_P - E_0} \approx \frac{E_0}{E_P} \qquad \text{(S42)}$$

Within this $4{\times}4$ $\vec{k}\cdot\vec{p}$-model without spin-orbit splitting, the electron and light hole masses (S40) and (S42) are nearly identical.

Using the light hole reduced mass

$$\frac{1}{\mu_{lh}} = \frac{1}{m_{lh}^*} + \frac{1}{m_e^*} = \frac{E_P + E_0}{E_0} + \frac{E_P - E_0}{E_0} = \frac{2E_P}{E_0} \qquad \text{(S43)}$$

we can write the light hole and electron energies in more compact form

$$E_{3,4} = \frac{\hbar^2 k^2}{2m_0} + \frac{E_0}{2}\left(1 \pm \sqrt{1 + \frac{\hbar^2 k^2}{2m_0}\frac{2}{\mu_{lh} E_0}}\right). \qquad \text{(S44)}$$

It is interesting that the expressions (S44) and (S69) are the same in the small and large spin-orbit splitting approximations.

The nonparabolicity of the electron and light hole bands is contained in the square-root term in Eq. (S36), if we include higher-order terms in k^2. By retaining the next-order terms, we obtain

$$\begin{aligned} E_3 &\approx E_0 + \frac{\hbar^2 k^2}{2m_0}\left(1 + \frac{E_P}{E_0} - \frac{\hbar^2 k^2}{2m_0}\frac{E_P^2}{E_0^3} + 2\frac{\hbar^4 k^4}{4m_0^2}\frac{E_P^3}{E_0^5}\right), \\ E_4 &\approx \frac{\hbar^2 k^2}{2m_0}\left(1 - \frac{E_P}{E_0} + \frac{\hbar^2 k^2}{2m_0}\frac{E_P^2}{E_0^3} - 2\frac{\hbar^4 k^4}{4m_0^2}\frac{E_P^3}{E_0^5}\right). \qquad \text{(S45)} \end{aligned}$$

Since E_P is nearly the same for many common semiconductors, we recognize that nonparabolicity corrections are larger for semiconductors with smaller band gaps.

We can also start the Taylor expansion of the square root with Eq. (S44):

$$\begin{aligned} \tilde{E}_{3,4} &= \frac{E_0}{2}\left(1 \pm \sqrt{1 + \frac{\hbar^2 k^2}{2m_0}\frac{2}{\mu_{lh} E_0}}\right) = \\ &= \frac{E_0}{2}\left[1 \pm 1 \pm \frac{\hbar^2 k^2}{2m_0}\frac{1}{\mu_{lh} E_0} \mp \frac{1}{8}\left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{4}{\mu_{lh}^2 E_0^2} \pm \right. \\ &\quad \left. \pm \frac{1}{16}\left(\frac{\hbar^2 k^2}{2m_0}\right)^3 \frac{8}{\mu_{lh}^3 E_0^3}\right]. \qquad \text{(S46)} \end{aligned}$$

The light hole energy is the expression with the minus sign

$$\begin{aligned} E_4 &= -\frac{\hbar^2 k^2}{2m_0}\left[\left(\frac{1}{2\mu_{lh}} - 1\right) - \frac{\hbar^2 k^2}{2m_0}\frac{1}{4\mu_{lh}^2 E_0} + \right. \\ &\quad \left. + \left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{1}{4\mu_{lh}^3 E_0^2}\right] \qquad \text{(S47)} \end{aligned}$$

We note that the parabolic factor

$$\frac{1}{2\mu_{lh}} - 1 = \frac{E_P}{E_0} - 1 = \frac{E_P - E_0}{E_0} = \frac{1}{m_{lh}^*} \qquad \text{(S48)}$$

is the inverse light hole mass in the small spin-orbit splitting approximation. The next two higher-order terms in Eq. (S47) agree with those given by Eq. (S45), but written in terms of the reduced light hole mass. This makes the light hole energy

$$E_4 = -\frac{\hbar^2 k^2}{2m_0 m_{lh}^*}\left[1 - \frac{\hbar^2 k^2}{2m_0}\frac{m_{lh}^*}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{m_{lh}^*}{4\mu_{lh}^3 E_0^2}\right]. \qquad \text{(S49)}$$

Similarly, the conduction band energy is given by

$$\begin{aligned} E_3 &= E_0 + \frac{\hbar^2 k^2}{2m_0}\left[\left(\frac{1}{2\mu_{lh}}+1\right) - \frac{\hbar^2 k^2}{2m_0}\frac{1}{4\mu_{lh}^2 E_0} + \right. \\ &\quad \left. + \left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{1}{4\mu_{lh}^3 E_0^2}\right] \\ &= E_0 + \frac{\hbar^2 k^2}{2m_0 m_e^*} \times \\ &\times \left[1 - \frac{\hbar^2 k^2}{2m_0}\frac{m_e^*}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{m_e^*}{4\mu_{lh}^3 E_0^2}\right]. \end{aligned} \quad \text{(S50)}$$

The expression

$$\frac{1}{2\mu_{lh}} + 1 = \frac{E_P}{E_0} + 1 = \frac{1}{m_e^*} \quad \text{(S51)}$$

is the electron mass.

Figures S1 and S2 show the valence and conduction bands of GaAs calculated from a 4×4 $\vec{k}\cdot\vec{p}$-model without spin orbit-splitting as given by Eqs. (S35) and (S36) and in various approximations.

We need to add a word of caution about the convergence of the series given by Eqs. (S45): For this series to converge, the third term in the series should be smaller than the second one, and the fourth term smaller than the third one. For GaAs, $E_P/E_0 \approx 17$. The approximate convergence criterion is therefore

$$\frac{\hbar^2 k^2}{2m_0} < \frac{E_0}{E_P} E_0 = \frac{E_0}{17} = 0.088 \text{ eV} = 0.003 \text{ Ha}, \quad \text{(S52)}$$

where 1 Ha=2 Ry=27.2 eV is the atomic unit of energy. This implies

$$k^2 < 0.003 \quad \text{or} \quad k < 0.06 \quad \text{(S53)}$$

in atomic units, where k is measured in inverse Bohr radii. Inspection of Figs. S1 and S2 indeed shows that the approximations (S45) fail for k>0.06 or k^2>0.004. For larger wave vectors, the parabolic approximations (S37) and (S38) are actually closer to the exact 4×4 solution (S36) than the nonparabolic corrections (S45). In energy units, the nonparabolic corrections (S45) are only beneficial up to about 0.4 eV above the band gap.

S6. BAND ENERGIES FROM 8×8 $\vec{k}\cdot\vec{p}$-MODEL WITH SPIN-ORBIT COUPLING

We now proceed to discuss how a finite spin-orbit splitting Δ_0 affects the band structure in a 4×4 $\vec{k}\cdot\vec{p}$ model. Following Kane (1957), we use a basis set $|S\rangle$, $|X \pm iY\rangle/\sqrt{2}$, and $|Z\rangle$ and assume that the $\vec{k}$ vector is along the z-direction. For other directions, one can rotate the coordinate system using Eqs. (6-9) in Kane 1957. The matrix is doubled to include the two spin states. In addition to the $\vec{k}\cdot\vec{p}$ perturbation Hamiltonian $H_{\vec{k}}$, we also include the k-independent spin-orbit interaction as a second perturbation. This spin-orbit Hamiltonian is given in Eq. (1) of Kane (1957) and also by Yu & Cardona Eq. (2.45a)

$$H_{\rm SO} = \frac{\hbar}{4c^2 m_0^2}\left(\vec{\nabla}V \times \vec{p}\right)\cdot\vec{\sigma}, \quad \text{(S54)}$$

where $\vec{\sigma}$ is a vector formed by the three Pauli matrices and c is the speed of light. There is also a k-dependent spin-orbit interaction, which is neglected. Δ_0 is the matrix element of the spin-orbit Hamiltonian (Kane 1957)

$$\Delta_0 = \frac{3\hbar i}{4m_0^2 c^2}\left\langle X \left| \frac{\partial V}{\partial x}p_y - \frac{\partial V}{\partial y}p_x \right| Y \right\rangle. \quad \text{(S55)}$$

The resulting 8×8 $\vec{k}\cdot\vec{p}$-matrix consists of two identical 4×4 blocks of the form

$$H_0 + \tilde{H}_{\vec{k}} + H_{\rm SO} = \begin{pmatrix} \tilde{H}_4 & 0 \\ 0 & \tilde{H}_4 \end{pmatrix}, \quad \text{(S56)}$$

where the 4×4 block is given by, see Kane (1957) Eq. (4) or Kane (1966) Eq. (36)

$$\tilde{H}_{\vec{k}} = \begin{pmatrix} E_0 & 0 & -\frac{\hbar k}{m_0} iP & 0 \\ 0 & -\frac{2\Delta_0}{3} & \frac{\sqrt{2}\Delta_0}{3} & 0 \\ \frac{\hbar k}{m_0} iP & \frac{\sqrt{2}\Delta_0}{3} & -\frac{\Delta_0}{3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad \text{(S57)}$$

The tilde over $H_{\vec{k}}$ indicates that the free-electron kinetic energy has been subtracted as indicated in Eq. (S17). A different Hamiltonian (with the energy eigenvalues at k=0 on the diagonal) is obtained if one chooses the wave functions given by Eqs. (2.48-2.49) by Yu and Cardona.

The matrix (S57) has one obvious eigenvalue $\tilde{E}_1$=0 (which we can call the heavy hole) and the other three are determined from the characteristic polynomial of the upper 3×3 block

$$\begin{aligned} &\left(E_0 - \tilde{E}\right)\left(-\frac{2\Delta_0}{3} - \tilde{E}\right)\left(-\frac{\Delta_0}{3} - \tilde{E}\right) - \\ &-\frac{\hbar^2 k^2}{m_0^2}P^2\left(-\frac{2\Delta_0}{3} - \tilde{E}\right) - \frac{2}{9}\Delta_0^2\left(E_0 - \tilde{E}\right) = 0. \end{aligned} \quad \text{(S58)}$$

The first and third term in Eq. (S58) can be combined to cancel the quadratic spin-orbit splitting term. By changing the overall sign, the characteristic equation becomes a cubic equation of the form

$$\tilde{E}\left(\tilde{E} - E_0\right)\left(\tilde{E} + \Delta_0\right) - \frac{\hbar^2 k^2 E_P}{2m_0}\left(\tilde{E} + \frac{2\Delta_0}{3}\right) = 0, \quad \text{(S59)}$$

which was given by Kane (1957) in Eq. (10) and by Kane (1966) in Eq. (38). See also Bartoli (PRB 27, 2248, 1983) in more common recent notation identical to Eq. (S59). This becomes equal to the earlier characteristic equation (S33) if we set the spin-orbit splitting Δ_0=0. Kane (1966) calls Eq. (S59) the small-gap approximation, because it

ignores the interactions with other "remote" bands, for example the p-antibonding upper conduction band, the s-bonding lower valence band, or d-orbitals.

This characteristic equation (S59) is a cubic equation, which can be solved analytically with modern computers. We will show the results below, but first we discuss the approximation that the spin-orbit splitting is much larger than the band gap, which is appropriate for InSb.

S7. APPROXIMATION FOR LARGE SPIN-ORBIT SPLITTING

If the spin-orbit splitting Δ_0 is large, then we can make the approximation $\tilde{E}+\Delta_0 \approx \Delta_0$. The secular equation (S59) then becomes quadratic and reduces to

$$\tilde{E}\left(\tilde{E}-E_0\right)-\frac{2}{3}\frac{\hbar^2k^2}{2m_0}E_P=0 \qquad \text{(S60)}$$

because the Δ_0 term can be cancelled. This characteristic equation is the same as in Eq. (S33) except for a factor $\frac{2}{3}$. The conduction and light hole bands are then described by an expression similar to Eq. (S36), except for the additional factor $\frac{2}{3}$, and are given by

$$\begin{aligned} E_{3,4} &= \frac{\hbar^2k^2}{2m_0}+\frac{E_0}{2}\left(1\pm\sqrt{1+\frac{8\hbar^2k^2P^2}{3m_0^2E_0^2}}\right)= \\ &= \frac{\hbar^2k^2}{2m_0}+\frac{E_0}{2}\left(1\pm\sqrt{1+\frac{8}{3}\frac{\hbar^2k^2E_P}{2m_0E_0^2}}\right), \end{aligned} \qquad \text{(S61)}$$

compare Eq. (13) in Kane (1957) or Eq. (50) in Kane (1966). The heavy hole band in the 8×8 approximation (which ignores interactions with the p-antibonding conduction band) still has the wrong curvature given by Eq. (S35). The split-off hole band E_2 is not included in this approximation.

We can expand the square root for small values of k, which leads to the parabolic approximation

$$\begin{aligned} E_3 &= E_0+\frac{\hbar^2k^2}{2m_0}+\frac{2}{3}\frac{\hbar^2k^2P^2}{m_0^2E_0}= \\ &= E_0+\frac{\hbar^2k^2}{2m_0}\left(\frac{2}{3}\frac{E_P}{E_0}+1\right), \end{aligned} \qquad \text{(S62)}$$

$$E_4 = \frac{\hbar^2k^2}{2m_0}-\frac{2}{3}\frac{\hbar^2k^2P^2}{m_0^2E_0}=-\frac{\hbar^2k^2}{2m_0}\left(\frac{2}{3}\frac{E_P}{E_0}-1\right) \qquad \text{(S63)}$$

The effective electron mass in this approximation for large spin-orbit splitting is

$$\frac{1}{m_e^*}=1+\frac{2}{3}\frac{E_P}{E_0}=\frac{2E_P+3E_0}{3E_0}\approx\frac{2}{3}\frac{E_P}{E_0} \qquad \text{(S64)}$$

or

$$m_e^*=\frac{3E_0}{2E_P+3E_0}\approx\frac{3}{2}\frac{E_0}{E_P}. \qquad \text{(S65)}$$

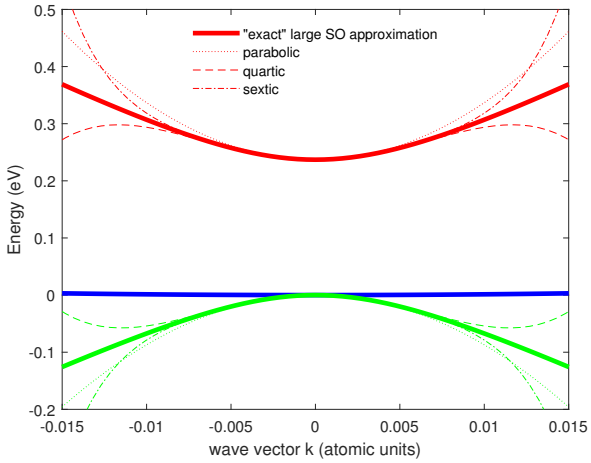

FIG. S3. Band structure of InSb near the Brillouin zone center calculated from a 4×4 $\vec{k}\cdot\vec{p}$-model within the approximation for large spin-orbit splitting. The electron, light hole, and heavy hole bands are shown in red, green, and blue, respectively. Thick lines show the dispersion calculated from Eqs. (S61) and dotted lines the parabolic approximation (S62) and (S63). Dashed and dot-dashed lines include the corrections proportional to k^4 and k^6 given by Eq. (S70) and (S71). The wave vector is shown in atomic units (inverse Bohr radii).

as expected from Eq. (S31) with a large spin-orbit splitting Δ_0. Similary, the light hole mass for large spin-orbit splittings is

$$\frac{1}{m_{lh}^*}=-\left(1-\frac{2}{3}\frac{E_P}{E_0}\right)=\frac{2E_P-3E_0}{3E_0}\approx\frac{2}{3}\frac{E_P}{E_0} \qquad \text{(S66)}$$

or

$$m_{lh}^*=\frac{3E_0}{2E_P-3E_0}\approx\frac{3}{2}\frac{E_0}{E_P}, \qquad \text{(S67)}$$

if we keep in mind that the overall minus sign is needed to ensure a negative curvature for a positive hole mass. This also agrees with Eq. (S31) in the limit of large Δ_0. The electron and light hole masses are again nearly identical for $E_0 \ll E_P$.

Using the light hole reduced mass

$$\frac{1}{\mu_{lh}}=\frac{1}{m_{lh}}+\frac{1}{m_e}=\frac{2E_P-3E_0}{3E_0}+\frac{2E_P+3E_0}{3E_0}=\frac{4E_P}{3E_0}, \qquad \text{(S68)}$$

we can write the light hole and electron energies in more compact form

$$E_{3,4}=\frac{\hbar^2k^2}{2m_0}+\frac{E_0}{2}\left(1\pm\sqrt{1+\frac{\hbar^2k^2}{2m_0}\frac{2}{\mu_{lh}E_0}}\right). \qquad \text{(S69)}$$

It is interesting that the expressions (S44) and (S69) are the same in the small and large spin-orbit splitting approximations.

The nonparabolicity of the electron and light hole bands is contained in the square-root term in Eq. (S61),

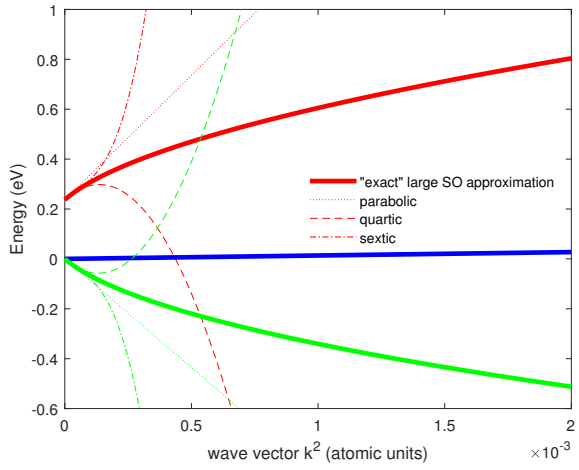

FIG. S4. As Fig. S3, but as a function of the square of the wave vector in atomic units. For comparison, with Fig. 1 in Kane (1957), the range of wave vectors was expanded.

if we include higher-order terms in k^2. By retaining the next-order terms, we obtain

$$E_3 = E_0 + \frac{\hbar^2 k^2}{2m_0} \quad \times \\ \times \left(1 + \frac{2}{3}\frac{E_P}{E_0} - \frac{4}{9}\frac{\hbar^2 k^2}{2m_0}\frac{E_P^2}{E_0^3} + 2\frac{8}{27}\frac{\hbar^4 k^4}{4m_0^2}\frac{E_P^3}{E_0^5}\right), \quad \text{(S70)}$$

$$E_4 = \frac{\hbar^2 k^2}{2m_0} \quad \times \\ \times \left(1 - \frac{2}{3}\frac{E_P}{E_0} + \frac{4}{9}\frac{\hbar^2 k^2}{2m_0}\frac{E_P^2}{E_0^3} - 2\frac{8}{27}\frac{\hbar^4 k^4}{4m_0^2}\frac{E_P^3}{E_0^5}\right). \quad \text{(S71)}$$

These equations are similar to those for small spin-orbit splittings given by Eqs. (S45) except for the powers of $\frac{2}{3}$ that appear as factors.

We can also start the Taylor expansion of the square root with Eq. (S69):

$$\begin{aligned}\tilde{E}_{3,4} &= \frac{E_0}{2}\left(1 \pm \sqrt{1 + \frac{\hbar^2 k^2}{2m_0}\frac{2}{\mu_{lh}E_0}}\right) = \\ &= \frac{E_0}{2}\left[1 \pm 1 \pm \frac{\hbar^2 k^2}{2m_0}\frac{1}{\mu_{lh}E_0} \mp \frac{1}{8}\left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{4}{\mu_{lh}^2 E_0^2} \pm \right. \\ &\left. \pm \frac{1}{16}\left(\frac{\hbar^2 k^2}{2m_0}\right)^3 \frac{8}{\mu_{lh}^3 E_0^3}\right]. \end{aligned} \quad \text{(S72)}$$

The light hole energy is the expression with the minus sign

$$\begin{aligned}E_4 &= -\frac{\hbar^2 k^2}{2m_0}\left[\left(\frac{1}{2\mu_{lh}} - 1\right) - \frac{\hbar^2 k^2}{2m_0}\frac{1}{4\mu_{lh}^2 E_0} + \right. \\ &\left. + \left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{1}{4\mu_{lh}^3 E_0^2}\right] \end{aligned} \quad \text{(S73)}$$

We note that the parabolic factor

$$\frac{1}{2\mu_{lh}} - 1 = \frac{2E_P}{3E_0} - 1 = \frac{2E_P - 3E_0}{3E_0} = \frac{1}{m_{lh}^*} \quad \text{(S74)}$$

is the inverse light hole mass in the large spin-orbit splitting approximation. The next two higher-order terms in Eq. (S47) agree with those given by Eq. (S45), but written in terms of the reduced light hole mass. This makes the light hole energy

$$E_4 = -\frac{\hbar^2 k^2}{2m_0 m_{lh}^*}\left[1 - \frac{\hbar^2 k^2}{2m_0}\frac{m_{lh}^*}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{m_{lh}^*}{4\mu_{lh}^3 E_0^2}\right], \quad \text{(S75)}$$

just like Eq. (S49) in the case of small spin-orbit splittings.

Similarly, the conduction band energy is given by

$$\begin{aligned}E_3 &= E_0 + \frac{\hbar^2 k^2}{2m_0}\left[\left(\frac{1}{2\mu_{lh}} + 1\right) - \frac{\hbar^2 k^2}{2m_0}\frac{1}{4\mu_{lh}^2 E_0} + \right. \\ &\left. + \left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{1}{4\mu_{lh}^3 E_0^2}\right] \\ &= E_0 + \frac{\hbar^2 k^2}{2m_0 m_e^*} \times \\ &\times \left[1 - \frac{\hbar^2 k^2}{2m_0}\frac{m_e^*}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{m_e^*}{4\mu_{lh}^3 E_0^2}\right], \end{aligned} \quad \text{(S76)}$$

just like Eq. (S47) in the case of small spin-orbit splittings. The expression

$$\frac{1}{2\mu_{lh}} + 1 = \frac{2E_P}{3E_0} + 1 = \frac{1}{m_e^*} \quad \text{(S77)}$$

is the electron mass, just like Eq. (S51) in the small spin-orbit splitting approximation.

Kane (1966) Eq. (50) and (1957) Eq. (13) also lists the dispersion of the split-off hole

$$\begin{aligned}E_2 &= -\Delta_0 + \frac{\hbar^2 k^2}{2m_0} - \frac{\hbar^2 k^2 P^2}{3m_0^2 (E_0 + \Delta_0)} = \\ &= -\Delta_0 + \frac{\hbar^2 k^2}{2m_0} - \frac{\hbar^2 k^2}{2m_0}\frac{E_P}{3(E_0 + \Delta_0)} \\ &= -\Delta_0 - \frac{\hbar^2 k^2}{2m_0}\left[\frac{E_P}{3(E_0 + \Delta_0)} - 1\right] \end{aligned} \quad \text{(S78)}$$

See also Yu & Cardona Eq. (2.59). We will derive this expression later in the context of the perturbative solution to the cubic characteristic equation in Sec. S9. The inverse split-off mass if therefore

$$\frac{1}{m_{\rm so}^*} = \frac{E_P}{3(E_0 + \Delta_0)} - 1. \quad \text{(S79)}$$

In principle, these expressions are expected to be quite accurate for InSb, where the condition $\Delta_0 \gg E_0$ is satisfied, except for the contributions of "remote" bands.

(For α-Sn, on the other hand, $\bar{E}_0$ is about half of Δ_0 and therefore the approximation for large spin-orbit splitting will be less accurate.) However, since $E_P/E_0 \approx 100$ for InSb, the range of applicability of Eqs. (S70) and (S71) is even more limited than for GaAs. Using the same convergence criterion (S52) as for GaAs, we expect the nonparabolicity corrections to converge only if

$$k^2 < 0.1 \times 10^{-3} \quad \text{or} \quad k < 0.01 \tag{S80}$$

for InSb (in atomic units for k). In energy units, the nonparabolic corrections (S70) and (S71) are not useful more than 70 meV from the band edge.

Figure S3 shows the top valence bands and the lowest conduction band of InSb in the center of the Brillouin zone. The dotted lines show the parabolic electron and light hole bands. The electron and light hole masses were calculated with Eq. (S31) using the parameters E_0=0.237 eV, Δ_0=0.81 eV, and E_P=23.1 eV given by Lawaetz 1971. The thick solid lines show the nonparabolic bands for the electron, heavy hole and light hole bands calculated using Eqs. (S61) within the large spin-orbit splitting approximation. The nonparabolic corrections given by Eqs. (S70) and (S71) are only applicable for very small wave vectors, as shown by the dashed and dot-dashed lines.

For comparison with Fig. 1 in Kane (1957), we also show the energies near k=0 on an expanded wave vector scale as a function of the square of the wave vector in atomic units, see Fig. S4. Despite neglecting the influence of "remote" bands, the agreement with Kane (1957) is quite good.

Additional improvements to the electron and hole dispersions for large spin-orbit splittings were given by Kane (1966) in Eqs. (52-56). Also, corrections to the electron and hole dispersions were given by Earnest Johnson (Semiconductors and Semimetals, Vol. 3, 1967), Eqs. (41-44) without an explanation.

S8. VIETA'S SOLUTION TO THE CUBIC CHARACTERISTIC EQUATION

The eigenvalues of the 4×4 Hamiltonian (S57) are given by the characteristic equation (S59)

$$\tilde{E}\left(\tilde{E}-E_0\right)\left(\tilde{E}+\Delta_0\right)-\frac{\hbar^2k^2}{2m_0}E_P\left(\tilde{E}+\frac{2\Delta_0}{3}\right)=0. \tag{S81}$$

With the substitution

$$A^2=\frac{1}{3}\frac{\hbar^2k^2}{2m_0}E_P, \tag{S82}$$

this can also be written as

$$\tilde{E}\left(\tilde{E}-E_0\right)\left(\tilde{E}+\Delta_0\right)-3A^2\left(\tilde{E}+\frac{2\Delta_0}{3}\right)=0. \tag{S83}$$

We collect the powers of $\tilde{E}$ to find the cubic equation

$$\tilde{E}^3+\tilde{E}^2\left(\Delta_0-E_0\right)-\tilde{E}\left(E_0\Delta_0+3A^2\right)-2A^2\Delta_0=0. \tag{S84}$$

This equation is of the form

$$ax^3+bx^2+cx+d=0 \tag{S85}$$

with the coefficients

$$\begin{aligned} x &= \tilde{E}=E+\frac{\hbar^2k^2}{2m_0}, && \text{(S86)}\\ a &= 1, && \text{(S87)}\\ b &= \Delta_0-E_0, && \text{(S88)}\\ c &= -E_0\Delta_0-3A^2, && \text{(S89)}\\ d &= -2A^2\Delta_0. && \text{(S90)} \end{aligned}$$

This general cubic equation can be reduced to a depressed cubic equation (see Wikipedia)

$$t^3+pt+q=0 \tag{S91}$$

with the substitution

$$\begin{aligned} t &= x+\frac{b}{3a}=x+\frac{1}{3}\left(\Delta_0-E_0\right), && \text{(S92)}\\ p &= \frac{3ac-b^2}{3a^2}= \\ &= -E_0\Delta_0-3A^2-\frac{1}{3}\left(\Delta_0-E_0\right)^2, && \text{(S93)}\\ q &= \frac{2b^3-9abc+27a^2d}{27a^3}=\frac{2}{27}\left(\Delta_0-E_0\right)^3+ \\ &+ \frac{1}{3}\left(\Delta_0-E_0\right)\left(E_0\Delta_0+3A^2\right)-2A^2\Delta_0. && \text{(S94)} \end{aligned}$$

The three solutions to the depressed cubic equation (due to Vieta, see Wikipedia) are given by

$$t_n=2\sqrt{\frac{-p}{3}}\cos\left[\frac{1}{3}\arccos\left(\frac{3q}{2p}\sqrt{\frac{-3}{p}}\right)-\frac{2\pi n}{3}\right] \tag{S95}$$

for n=0, 1, 2. We note that p is negative. Therefore, the arguments under the square roots in Eq. (S95) are real.

Using these expressions, the exact solution to the cubic characteristic equation for the 4×4 $\vec{k}\cdot\vec{p}$ Hamiltonian can be programmed in MATLAB. Results are shown in Figs. S5 and S6 for GaAs and InSb together with the appropriate small or large spin-orbit approximation. The corrections for the conduction band are moderate, but large for the light hole band in GaAs, which strongly couples to the split-off hole band, as mentioned by Bartoli (1983).

Figure S7 shows the inverted band structure of α-tin with the Γ_7^- s-antibonding valence band in green and the p-bonding Γ_8^+ conduction band in red, calculated from the 4×4 $\vec{k}\cdot\vec{p}$ Kane Hamiltonian (S57) (solid). The deviations from the parabolic bands (dotted) are even stronger for α-tin than for GaAs and InSb, presumably because of

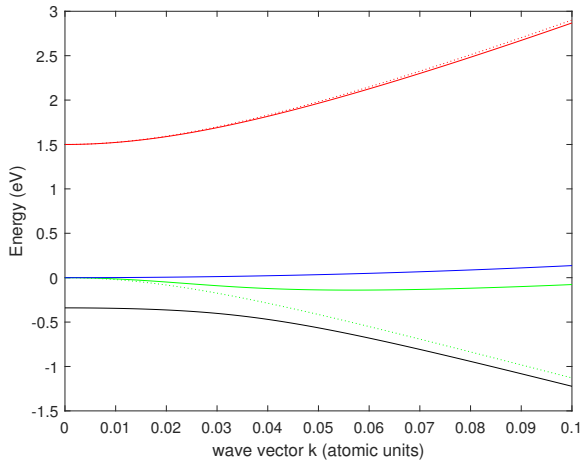

FIG. S5. Heavy (blue), light (green), and split-off hole (black) and electron bands (red) of GaAs near k=0 calculated from the cubic characteristic equation (S59) of the 4×4 $\vec{k}\cdot\vec{p}$ Kane Hamiltonian (S57) for finite spin-orbit splitting (solid) and from the small spin-orbit approximation (S36) (dotted).

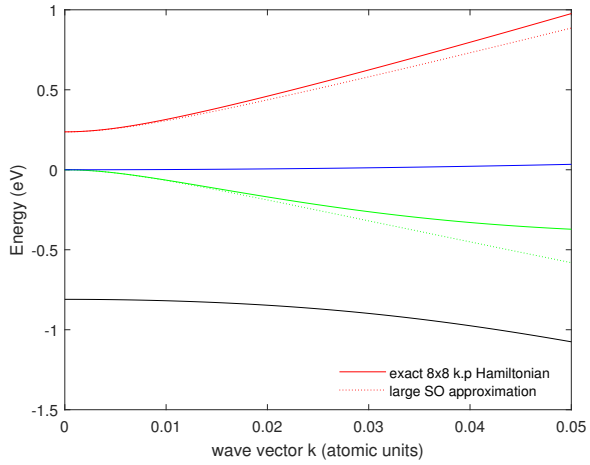

FIG. S6. Heavy (blue), light (green), and split-off hole (black) and electron bands (red) of InSb near k=0 calculated from the cubic characteristic equation (S59) of the 4×4 $\vec{k}\cdot\vec{p}$ Kane Hamiltonian (S57) for finite spin-orbit splitting (solid) and from the large spin-orbit approximation (S61) (dotted).

the strong interaction of these bands with the Γ_7^+ split-off hole band. The Γ_7^- and Γ_8^+ valence bands are nearly parallel across a large volume of the Brillouin zone, giving rise to strong allowed intravalence band transitions at the inverted E_0 band gap in p-type α-tin or at room temperature.

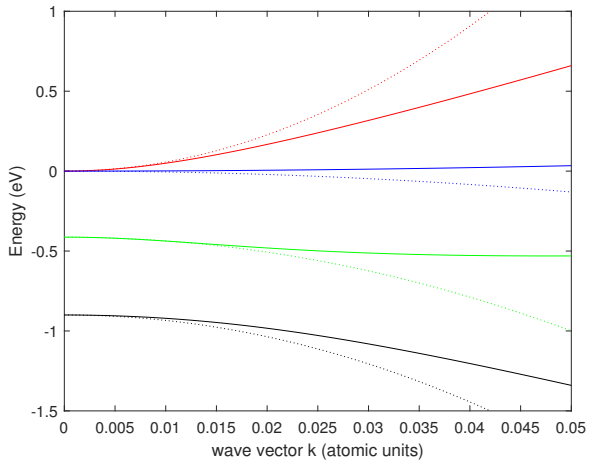

FIG. S7. Γ_8^+ conduction (red) and valence (blue) bands, Γ_7^- valence band (green), and Γ_7^+ split-off band for α-Sn calculated from the 4×4 $\vec{k}\cdot\vec{p}$ Kane Hamiltonian (S57) (solid) and the parabolic bands (dotted) with effective masses given by Carrasco (2018).

S9. PERTURBATIVE SOLUTION OF THE 4×4 CHARACTERISTIC EQUATION

The approximations for small and large spin-orbit splittings derived in Secs. S5 and S6 resulted in expressions (S40) and (S65) for the effective electron mass. Neither of these are satisfactory, because they do not yield the electron mass (S31) calculated using non-degenerate perturbation theory. Instead, we pursue a perturbative approach to solve the cubic characteristic equation

$$\tilde{E}^3 + \tilde{E}^2\left(\Delta_0 - E_0\right) - \tilde{E}\left(E_0\Delta_0 + 3A^2\right) - 2A^2\Delta_0 = 0 \tag{S96}$$

for small wave vectors (but arbitrary values of E_0 and Δ_0), where A was defined in Eq. (S82). With the substitution

$$\epsilon = \frac{3A^2}{E_0\Delta_0} = \frac{\hbar^2k^2}{2m_0}\frac{E_P}{\Delta_0 E_0}, \tag{S97}$$

the characteristic equation becomes

$$\tilde{E}^3 + \tilde{E}^2\left(\Delta_0 - E_0\right) - \tilde{E}E_0\Delta_0\left(1+\epsilon\right) - \frac{2}{3}\epsilon E_0\Delta_0^2 = 0, \tag{S98}$$

For small wave vectors (and thus small values of ϵ), we write the modified energy $\tilde{E}$ as a Taylor series in powers of ϵ, up to second order

$$\tilde{E} = \lambda_0 + \lambda_1\epsilon + \lambda_2\epsilon^2 \tag{S99}$$

and substitute Eq. (S99) into Eq. (S98). With the intermediate expressions

$$\tilde{E}^2 \approx \lambda_0^2 + 2\lambda_0\lambda_1\epsilon + \left(\lambda_1^2 + 2\lambda_0\lambda_2\right)\epsilon^2, \tag{S100}$$

$$\tilde{E}^3 \approx \lambda_0^3 + 3\lambda_0^2\lambda_1\epsilon + 3\left(\lambda_0^2\lambda_2 + \lambda_0\lambda_1^2\right)\epsilon^2, \tag{S101}$$

the characteristic equation (S98) becomes

$$\begin{aligned}\lambda_0^3 + 3\lambda_0^2\lambda_1\epsilon + 3\left(\lambda_0^2\lambda_2 + \lambda_0\lambda_1^2\right)\epsilon^2 + \\ + \left[\lambda_0^2 + 2\lambda_0\lambda_1\epsilon + \left(\lambda_1^2 + 2\lambda_0\lambda 2\right)\epsilon^2\right](\Delta_0 - E_0) - \\ - \left(\lambda_0 + \lambda_1\epsilon + \lambda_2\epsilon^2\right)E_0\Delta_0(1+\epsilon) - \frac{2}{3}\epsilon E_0\Delta_0^2 = 0\end{aligned} \tag{S102}$$

This characteristic equation (S102) must be valid for all orders of ϵ. This yields three equations for the Taylor series coefficients in Eq. (S99).

$$\lambda_0^3 + \lambda_0^2(\Delta_0 - E_0) - \lambda_0 E_0\Delta_0 = 0, \tag{S103}$$

$$\begin{aligned}3\lambda_0^2\lambda_1 + 2\lambda_0\lambda_1(\Delta_0 - E_0) - \\ - (\lambda_0 + \lambda_1)E_0\Delta_0 - \frac{2}{3}E_0\Delta_0^2 = 0,\end{aligned} \tag{S104}$$

$$\begin{aligned}3\left(\lambda_0^2\lambda_2 + \lambda_0\lambda_1^2\right) + \\ + \left(\lambda_1^2 + 2\lambda_0\lambda_2\right)(\Delta_0 - E_0) - \\ - (\lambda_1 + \lambda_2)E_0\Delta_0 = 0.\end{aligned} \tag{S105}$$

If we are interested in the dispersion of the light hole bands, then we know that λ_0=0. Equation (S104) therefore yields

$$\lambda_1 = -\frac{2}{3}\Delta_0 \tag{S106}$$

and Eq. (S105) yields

$$\lambda_2 = \frac{4}{9}\frac{\Delta_0^2}{E_0} + \frac{2}{9}\Delta_0 = \frac{2}{9}\Delta_0\left(1 + \frac{2\Delta_0}{E_0}\right). \tag{S107}$$

We insert this into Eq. (S99) and find the modified light hole energy to be

$$\tilde{E}_{\rm lh} = -\frac{2}{3}\frac{\hbar^2k^2}{2m_0}\frac{E_P}{E_0} + \frac{2}{9}\left(1 + \frac{2\Delta_0}{E_0}\right)\left(\frac{\hbar^2k^2}{2m_0}\right)^2\frac{E_P^2}{\Delta_0 E_0^2}. \tag{S108}$$

The light hole energy $E_{\rm lh}$ is determined from Eq. (S17) and turns out to be

$$E_{\rm lh} = \frac{\hbar^2k^2}{2m_0}\left[1 - \frac{2}{3}\frac{E_P}{E_0} + \frac{2}{9}\frac{E_P^2}{\Delta_0 E_0^2}\left(1 + \frac{2\Delta_0}{E_0}\right)\frac{\hbar^2k^2}{2m_0}\right]. \tag{S109}$$

The first (quadratic) term in this expression yields the light hole mass

$$\frac{1}{m_{\rm lh}^*} = \frac{2}{3}\frac{E_P}{E_0} - 1 = \frac{2E_P - 3E_0}{3E_0}. \tag{S110}$$

This is the same expression as in Eq. (S66) for the large spin-orbit approximation. For the small spin-orbit splitting approximation, Eq. (S42) yields a different result, because the strong interaction with the split-off band has been neglected. The second (quartic) term in Eq. (S109) approaches the lowest-order nonparabolic correction in the limit $\Delta_0 \gg E_0$ given by Eq. (S71), but offers a modification for a finite spin-orbit splitting. As shown in Fig. S8, the perturbative solution (S109) to the cubic characteristic equation actually provides a poorer approximation to the full solution in Sec. S8 than the lowest-order nonparabolic correction in Eq. (S71) in the approximation for large spin-orbit spittings.

We find the dispersion of the split-off hole band by setting $\lambda_0 = -\Delta_0$. For this band, Eq. (S104) yields

$$\lambda_1 = -\frac{1}{3}\frac{E_0\Delta_0}{E_0 + \Delta_0} \tag{S111}$$

and Eq. (S105) yields

$$\lambda_2 = -\frac{E_0^2\Delta_0(\Delta_0 + 2E_0)}{9(E_0 + \Delta_0)^3}. \tag{S112}$$

We insert these expressions into Eq. (S99) and find the split-off hole energy

$$\begin{aligned}E_{\rm so} &= -\Delta_0 - \frac{\hbar^2k^2}{2m_0}\left[\frac{E_P}{3(E_0 + \Delta_0)} - 1\right] - \\ &- \left(\frac{\hbar^2k^2}{2m_0}\right)^2\frac{E_P^2}{\Delta_0}\frac{\Delta_0 + 2E_0}{9(E_0 + \Delta_0)^3}.\end{aligned} \tag{S113}$$

The quadratic term in Eq. (S113) contains the inverse effective mass of the split-off hole (in square brackets), which we have already mentioned in Eq. (S79).

Finally, we find the dispersion of the conduction band by setting $\lambda_0 = E_0$. For this band, Eq. (S104) yields

$$\lambda_1 = \frac{\Delta_0\left(E_0 + \frac{2}{3}\Delta_0\right)}{E_0 + \Delta_0} = \frac{E_0\Delta_0}{3}\left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0}\right) \tag{S114}$$

and the parabolic expression for the conduction band is therefore

$$E_e = E_0 + \frac{\hbar^2k^2}{2m_0}\left[1 + \frac{E_P}{E_0}\frac{E_0 + \frac{2}{3}\Delta_0}{E_0 + \Delta_0}\right] \tag{S115}$$

$$= E_0 + \frac{\hbar^2k^2}{2m_0}\left[1 + \frac{E_P}{3}\left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0}\right)\right] \tag{S116}$$

$$= E_0 + \frac{\hbar^2k^2}{2m_0m_e}. \tag{S117}$$

The expression in square brackets is the inverse effective electron mass

$$\frac{1}{m_e^*} = 1 + \frac{E_P}{3}\left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0}\right) \tag{S118}$$

given already in (S31) as found from non-degenerate perturbation theory. It also agrees with Eq. (S40) in the limit of Δ_0=0 and with Eq. (S65) in the limit of large spin-orbit splittings. The inverse reduced light hole mass in this approxmation is

$$\frac{1}{\mu_{lh}} = \frac{1}{m_{lh}^*} + \frac{1}{m_e^*} = \frac{E_P}{3}\left(\frac{4}{E_0} + \frac{1}{E_0 + \Delta_0}\right). \tag{S119}$$

Equation (S105) yields

$$\lambda_2 = -\frac{\lambda_1\left[\lambda_1(2E_0 + \Delta_0) - E_0\Delta_0\right]}{E_0(E_0 + \Delta_0)}. \tag{S120}$$

We note that

$$\lambda_1\epsilon = \left(\frac{1}{m_e} - 1\right)\frac{\hbar^2 k^2}{2m_0}. \tag{S121}$$

and therefore

$$\begin{aligned}\lambda_2\epsilon^2 &= -\frac{\lambda_1\epsilon\left[\lambda_1\epsilon\left(2E_0+\Delta_0\right) - E_0\Delta_0\epsilon\right]}{E_0\left(E_0+\Delta_0\right)} = \\ &= -\left(\frac{1}{m_e}-1\right)^2 \frac{2E_0+\Delta_0-\frac{E_P}{\frac{1}{m_e}-1}}{E_0\left(E_0+\Delta_0\right)}\left(\frac{\hbar^2k^2}{2m_0}\right)^2\end{aligned}$$

The middle factor can be simplified using Eq. (S31) as follows

$$\begin{aligned}&\frac{2E_0+\Delta_0-\frac{E_P}{\frac{1}{m_e}-1}}{E_0\left(E_0+\Delta_0\right)} = \\ &= \frac{2E_0+\Delta_0-\frac{E_P}{\frac{E_P}{3}\frac{3E_0+2\Delta_0}{E_0(E_0+\Delta_0)}}}{E_0\left(E_0+\Delta_0\right)} = \\ &= \frac{2E_0+\Delta_0-3\frac{E_0(E_0+\Delta_0)}{3E_0+2\Delta_0}}{E_0\left(E_0+\Delta_0\right)} = \\ &= \frac{3E_0^2+2E_0\Delta_0+2E_0\Delta_0+2\Delta_0^2}{E_0\left(E_0+\Delta_0\right)\left(3E_0+2\Delta_0\right)} = \\ &= \frac{E_0\left(3E_0+2\Delta_0\right)+2\Delta_0\left(E_0+\Delta_0\right)}{E_0\left(E_0+\Delta_0\right)\left(3E_0+2\Delta_0\right)} = \\ &= \frac{1}{E_0+\Delta_0}+\frac{2\Delta_0}{E_0}\frac{1}{3E_0+2\Delta_0}.\end{aligned} \tag{S122}$$

We insert these expressions into Eq. (S99) and find the conduction band electron energy

$$\begin{aligned}E_e &= E_0+\frac{\hbar^2k^2}{2m_0}\left[1+\frac{E_P}{E_0}\frac{E_0+\frac{2}{3}\Delta_0}{E_0+\Delta_0}\right] - \\ &- \left(\frac{1}{m_e}-1\right)^2\left(\frac{1}{E_0+\Delta_0}+\frac{2\Delta_0}{E_0}\frac{1}{3E_0+2\Delta_0}\right)\left(\frac{\hbar^2k^2}{2m_0}\right)^2 \\ &= E_0+\frac{\hbar^2k^2}{2m_0}\left[1+\frac{E_P}{3}\left(\frac{2}{E_0}+\frac{1}{E_0+\Delta_0}\right)\right] - \\ &- \left(\frac{1}{m_e}-1\right)^2\left(\frac{1}{E_0+\Delta_0}+\frac{2\Delta_0}{E_0}\frac{1}{3E_0+2\Delta_0}\right)\left(\frac{\hbar^2k^2}{2m_0}\right)^2\end{aligned} \tag{S123}$$

The result of this approximation (S123) is shown in Fig. S8. It provides about the same accuracy as the lowest-order nonparabolic correction for large spin-orbit splittings given by Eq. (S71).

S10. HEURISTIC INTERPOLATION BETWEEN SMALL AND LARGE SPIN-ORBIT SPLITTINGS

As we have seen previously during the derivation of the light hole and electron energies for small and large spin-orbit splittings, both approximations yield the same expressions given by Eqs. (S44) and (S69)

$$E_{3,4} = \frac{\hbar^2k^2}{2m_0}+\frac{E_0}{2}\left(1\pm\sqrt{1+\frac{\hbar^2k^2}{2m_0}\frac{2}{\mu_{lh}E_0}}\right), \tag{S124}$$

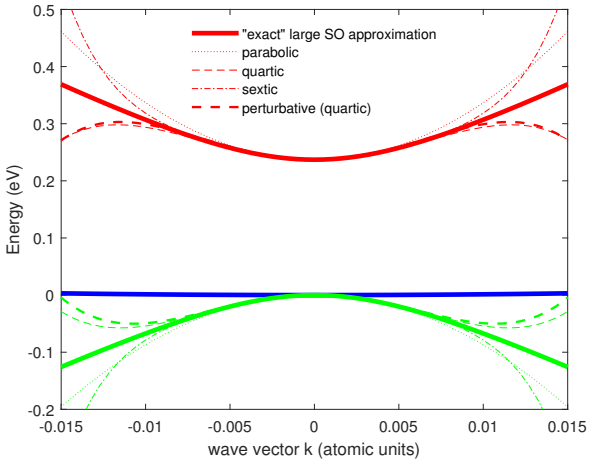

FIG. S8. As Fig. S6 for InSb, but we have added thick dashed lines to indicate the light hole and electron dispersions from the perturbative solutions to the cubic characteristic equation, as given by Eqs. (S109) and (S123).

but the definitions of the reduced effective mass is different in both cases. We can expand this equation for small wave vectors, keeping only terms up to k^6, and find

$$E_4 = -\frac{\hbar^2k^2}{2m_0m_{lh}^*}\left[1-\frac{\hbar^2k^2}{2m_0}\frac{m_{lh}^*}{4\mu_{lh}^2E_0}+\left(\frac{\hbar^2k^2}{2m_0}\right)^2\frac{m_{lh}^*}{4\mu_{lh}^3E_0^2}\right]. \tag{S125}$$

for the light hole band and

$$\begin{aligned}E_3 &= E_0+\frac{\hbar^2k^2}{2m_0m_e^*}\times \\ &\times\left[1-\frac{\hbar^2k^2}{2m_0}\frac{m_e^*}{4\mu_{lh}^2E_0}+\left(\frac{\hbar^2k^2}{2m_0}\right)^2\frac{m_e^*}{4\mu_{lh}^3E_0^2}\right]\end{aligned} \tag{S126}$$

for the electron band.

It is tempting to use these expressions, although they were derived as approximations for two special cases, also for the general case of band gaps and spin-orbit splittings, with the light hole, electron, and reduced effective masses obtained from the perturbative solution, as given by Eqs. (S110), (S118), and (S119). This provides an elegant method of interpolation between the small and large spin-orbit approximations. One might also use the experimental masses or those calculated from a larger $\vec{k}\cdot\vec{p}$ Hamiltonian for improved accuracy.

Whether this approach will work, however, is not guaranteed and needs to be checked by comparison with more precise calculations. It is particularly problematic to use the electron (or light hole) mass from one type of approximation and the reduced mass from another.

Results are shown in Figs. S9 and S10. Only the exact solutions (S95) to the cubic equation and the square-root expression (S124) in the approximation for large spin-orbit splittings yield reasonable results. The expansion of the square root up to terms linear in k^6 only results in

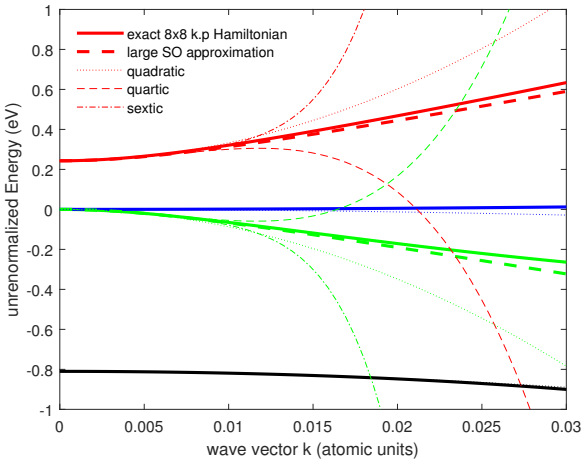

FIG. S9. Similar to Fig. S6 for InSb, but calculated with the reduced mass expressions (S124), (S125), and (S126). Thick lines show the heavy hole (blue), light hole (green), split-off hole (black), and electron bands (red) of the cubic characteristic equation given by Eq. (S95) (solid) and from the large spin-orbit approximation (S124) (dashed). Thin lines show the expansion of the square root including terms proportional to k^2 (dotted), k^4 (dashed), and k^6 (dot-dashed). Parabolic bands for the heavy and split off holes with experimental masses are also shown.

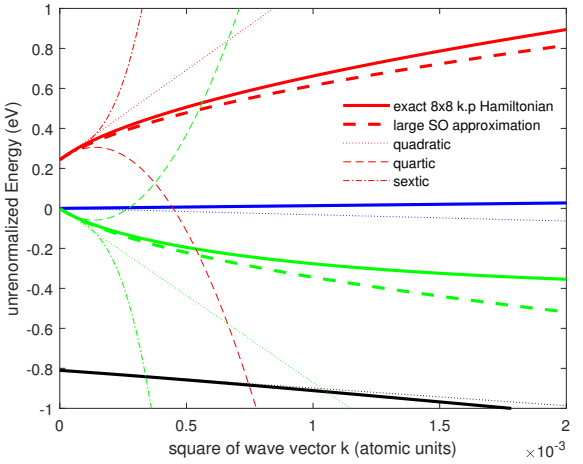

FIG. S10. Same as Fig. S9, but drawn with the square of the wave vector (in atomic units, i.e., inverse Bohr radii) on the horizontal axis.

a good approximation for a wave vector ka_B below 0.01 (energies of no more than 50 meV above the conduction band minimum). Due to the small band gap, nonparabolicity effects are very large. Because of the small gap, the expansion of the square root diverges very quickly and the coefficients become larger for higher-order terms.

S11. NONPARABOLICITY PARAMETERS

So far, our focus has been to find expressions for the band energies as a function of wave vector and we have plotted these results with various approximations derived from the 4×4 $\vec{k}\cdot\vec{p}$ Kane Hamiltonian (S57). Sometimes, it is also necessary to write the wave vector as a function of energy

$$\frac{\hbar^2k^2}{2m_0m_i}=\frac{K^2}{m_i} = \epsilon_i\left(1+\alpha_i\epsilon_i+\beta_i\epsilon_i^2\right) = \\ = \epsilon_i\left(1+\alpha_i'\frac{\epsilon_i}{E_0}+\beta_i'\frac{\epsilon_i^2}{E_0^2}\right), \quad \text{(S127)}$$

where i is the band index and the αs and βs are the lowest-order nonparabolicity coefficients. The units of the unprimed coefficients are an inverse power of energy, while the primed coefficients are dimensionless. For the heavy and light hole bands, we define $E_{\rm lh,hh}=-\epsilon_{\rm lh,hh}$ to keep the effective masses positive. For the electron and split-off hole bands, ϵ_i is the excess energy above or below the band extremum, i.e., $E_e=E_0+\epsilon_e$ for the conduction band and $E_{\rm so}=-\Delta_0-\epsilon_{\rm so}$ for the split-off hole band. To simplify the notation, we also define $K^2=\hbar^2k^2/2m_0$, which implies the use of atomic units (Bohr radius and Hartree as units for length and energy).

Since Vieta's solution (S95) to the cubic characteristic equation (S59) is "complicated" (see Bartoli 1983), we begin by deriving the nonparabolicity parameters in the small and large spin-orbit splitting approximations.

A. For small spin-orbit splittings

Conwell and Vassell (Phys. Rev. **166**, 797, 1968) introduce a non-parabolicity parameter α, which is required for the calculation of the joint density of states. Since they are interested primarily in GaAs, where $\Delta_0 \ll E_0$, they start with the characteristic equation for the conduction band for small spin-orbit splitting (S36)

$$\epsilon_e=K^2+\frac{E_0}{2}\left(\sqrt{1+4K^2\frac{E_P}{E_0^2}}-1\right), \quad \text{(S128)}$$

where ϵ_e is the energy referenced to the conduction band minimum at the Γ-point. This can be solved as a function of the square of the wave vector as follows:

$$\left(\epsilon_e-K^2+\frac{E_0}{2}\right)^2=\frac{E_0^2}{4}\left(1+4K^2\frac{E_P}{E_0^2}\right), \quad \text{(S129)}$$

$$K^4-K^2E_0-2K^2\epsilon_e+\epsilon_eE_0+\epsilon_e^2+\frac{1}{4}E_0^2=\frac{1}{4}E_0^2+K^2E_P, \quad \text{(S130)}$$

$$K^4-K^2\left(E_0+E_P+2\epsilon_e\right)+\epsilon_e\left(E_0+\epsilon_e\right)=0, \quad \text{(S131)}$$

This biquadratic equation has the solution

$$\begin{aligned} K^2 &= \frac{1}{2}\left[(E_0+E_P+2\epsilon_e) \pm \right. \\ &\pm \left. \sqrt{(E_0+E_P+2\epsilon_e)^2 - 4\epsilon_e\,(E_0+\epsilon_e)}\right] = \\ &= \epsilon_e + \frac{1}{2}\left[(E_0+E_P) \pm \sqrt{(E_0+E_P)^2+4\epsilon_e E_P}\right] = \\ &= \epsilon_e + \frac{E_0+E_P}{2}\left[1-\sqrt{1+\frac{4\epsilon_e E_P}{(E_0+E_P)^2}}\right]. \quad \text{(S132)} \end{aligned}$$

We have selected the solution with the minus sign, because we are looking for a solution where ϵ_e=0 for k=0. This equation (S132) is the same as Eq. (2.2) by Conwell and Vassell (1968).

For sufficiently small energies ϵ_e, we can expand the square root to find

$$\begin{aligned} K^2 &= \epsilon_e - \frac{\epsilon_e E_P}{E_0+E_P} + \frac{\epsilon_e^2 E_P^2}{(E_0+E_P)^3} - \frac{2\epsilon_e^3 E_P^3}{(E_0+E_P)^5} = \\ &= \epsilon_e\left[1 - \frac{E_P}{E_0+E_P} + \frac{\epsilon_e E_P^2}{(E_0+E_P)^3} - \frac{2\epsilon_e^2 E_P^3}{(E_0+E_P)^5}\right] = \\ &= \epsilon_e\left[\frac{E_0}{E_0+E_P} + \frac{\epsilon_e E_P^2}{(E_0+E_P)^3} - 2\frac{\epsilon_e^2 E_P^3}{(E_0+E_P)^5}\right] = \\ &= \epsilon_e\left[m_e^* + \frac{\epsilon_e E_P^2}{(E_0+E_P)^3} - 2\frac{\epsilon_e^2 E_P^3}{(E_0+E_P)^5}\right] = \\ &= m_e^*\epsilon_e\left[1 + \frac{\epsilon_e E_P^2}{m_e^*(E_0+E_P)^3} - 2\frac{\epsilon_e^2 E_P^3}{m_e^*(E_0+E_P)^5}\right] = \\ &= m_e^*\epsilon_e\left[1 + \frac{\epsilon_e E_P^2}{E_0(E_0+E_P)^2} - 2\frac{\epsilon_e^2 E_P^3}{E_0(E_0+E_P)^4}\right] \quad \text{(S133)} \end{aligned}$$

We have used Eq. (S40) for the effective electron mass within the small spin-orbit approximation. Equation (S133) is the same as Eq. (2.3) in Conwell and Vassell (1968).

We take the effective mass to the left hand side to arrive at the definition of the non-parabolicity parameters α_e and β_e for the conduction band

$$\begin{aligned} \frac{K^2}{m_e^*} &= \epsilon_e\left[1 + \frac{\epsilon_e E_P^2}{E_0(E_0+E_P)^2} - 2\frac{\epsilon_e^2 E_P^3}{E_0(E_0+E_P)^4}\right] = \\ &= \epsilon_e\left(1+\alpha_e\epsilon_e+\beta_e\epsilon_e^2\right) = \\ &= \epsilon_e\left(1+\alpha_e'\frac{\epsilon_e}{E_0}+\beta_e'\frac{\epsilon_e^2}{E_0^2}\right), \quad \text{(S134)} \end{aligned}$$

following the definition given by Eq. (2.4) in Conwell and Vassell (1968) or (for the primed quantities) the definition given by Bartoli (1983) in Eq. (A2). By comparing the coefficients, we find

$$\alpha_e = \frac{E_P^2}{E_0(E_0+E_P)^2} \approx \frac{1}{E_0}, \quad \text{(S135)}$$

$$\alpha_e' = \frac{E_P^2}{(E_0+E_P)^2} \approx 1, \quad \text{(S136)}$$

$$\beta_e = -\frac{2E_P^3}{E_0(E_0+E_P)^4} \approx -\frac{2}{E_0 E_P}, \quad \text{(S137)}$$

$$\beta_e' = -\frac{2E_0 E_P^3}{(E_0+E_P)^4} \approx -\frac{2E_0}{E_P}. \quad \text{(S138)}$$

These expressions were given by Conwell and Vassell (1968). The approximations are valid for $\epsilon \ll E_0$ and $E_0 \ll E_P$, which is usually the case. It seems that Eq. (S127) converges faster than Eq. (S45). Equations (S135-S138) agree with Eq. (A3a-A3b) in Bartoli (1983) for Δ_0=0.

We can apply the same formalism to the light hole band, where we start with Eq. (S36)

$$\begin{aligned} E_{lh} &= K^2 + \frac{E_0}{2}\left(1-\sqrt{1+4K^2\frac{E_P}{E_0^2}}\right) = \\ &= K^2 - \frac{E_0}{2}\left(\sqrt{1+4K^2\frac{E_P}{E_0^2}}-1\right). \quad \text{(S139)} \end{aligned}$$

This is the same equation as Eq. (S128) except that E_0 has the opposite sign. We also need to consider that the light hole energy E_{lh} is negative and its mass m_{lh} is positive. The outcome is therefore

$$\begin{aligned} K^2 &= E_{lh}\left[-\frac{E_0}{E_P-E_0} + \frac{E_{lh}E_P^2}{(E_P-E_0)^3} - 2\frac{E_{lh}^2 E_P^3}{(E_P-E_0)^5}\right] = \\ &= -E_{lh}\left[\frac{E_0}{E_P-E_0} - \frac{E_{lh}E_P^2}{(E_P-E_0)^3} + 2\frac{E_{lh}^2 E_P^3}{(E_P-E_0)^5}\right] = \\ &= -E_{lh}\left[m_{lh} - \frac{E_{lh}E_P^2}{(E_P-E_0)^3} + 2\frac{E_{lh}^2 E_P^3}{(E_P-E_0)^5}\right]. \quad \text{(S140)} \end{aligned}$$

By bringing the light hole effective mass over to the left hand side and using Eq. (S42), we obtain

$$\begin{aligned} \frac{K^2}{m_{lh}^*} &= -E_{lh}\left[1 - \frac{E_{lh}E_P^2}{E_0(E_P-E_0)^2} + 2\frac{E_{lh}^2 E_P^3}{E_0(E_P-E_0)^4}\right] = \\ &= -E_{lh}\left(1-\alpha_{lh}E_{lh}+\beta_{lh}E_{lh}^2\right) = \\ &= -E_{lh}\left(1-\alpha_{lh}'\frac{E_{lh}}{E_0}+\beta_{lh}'\frac{E_{lh}^2}{E_0^2}\right), \quad \text{(S141)} \end{aligned}$$

By comparing the coefficients, we find

$$\alpha_{lh} = \frac{E_P^2}{E_0(E_P-E_0)^2} \approx \frac{1}{E_0}, \quad \text{(S142)}$$

$$\alpha_{lh}' = \frac{E_P^2}{(E_P-E_0)^2} \approx 1, \quad \text{(S143)}$$

$$\beta_{lh} = \frac{2E_P^3}{E_0(E_P-E_0)^4} \approx \frac{2}{E_0 E_P}, \quad \text{(S144)}$$

$$\beta_{lh}' = \frac{2E_0 E_P^3}{(E_P-E_0)^4} \approx \frac{2E_0}{E_P}. \quad \text{(S145)}$$

Bartoli (1983) mentions that applying the small spin-orbit approximation to the light hole bands is not meaningful, since the light hole and split-off hole bands are strongly coupled for small spin orbit splittings. We have shown this already in Fig. S5.

B. For large spin-orbit splittings

Next, we discuss how the conduction band nonparabolicity parameters (S135-S138) need to be adjusted in the limit of large spin-orbit splittings. The characteristic equation (S61)

$$E_{3,4} = \frac{\hbar^2 k^2}{2m_0} + \frac{E_0}{2}\left(1 \pm \sqrt{1 + \frac{8}{3}\frac{\hbar^2 k^2 E_P}{2m_0 E_0^2}}\right) \tag{S146}$$

has an additional factor $\frac{2}{3}$ under the square root. We rewrite this equation for the conduction band as

$$\epsilon_e = K^2 + \frac{E_0}{2}\left(\sqrt{1 + \frac{8}{3}K^2\frac{E_P}{E_0^2}} - 1\right) \tag{S147}$$

to show the similarity with Eq. (S128) for vanishing spin-orbit splitting. We need to solve this equation for K^2, just like in the case for vanishing spin-orbit splitting.

$$\left(\epsilon_e - K^2 + \frac{E_0}{2}\right)^2 = \frac{E_0^2}{4}\left(1 + \frac{8}{3}K^2\frac{E_P}{E_0^2}\right), \tag{S148}$$

$$K^4 - K^2 E_0 - 2K^2\epsilon_e + \epsilon_e E_0 + \epsilon_e^2 + \frac{1}{4}E_0^2 = \frac{1}{4}E_0^2 + \frac{2}{3}K^2 E_P, \tag{S149}$$

$$K^4 - K^2\left(E_0 + \frac{2}{3}E_P + 2\epsilon_e\right) + \epsilon_e\left(E_0 + \epsilon_e\right) = 0, \tag{S150}$$

$$\begin{aligned} K^2 &= \frac{1}{2}\left[\left(E_0 + \frac{2}{3}E_P + 2\epsilon_e\right) \pm \right. \\ &\left. \pm \sqrt{\left(E_0 + \frac{2}{3}E_P + 2\epsilon_e\right)^2 - 4\epsilon_e\left(E_0 + \epsilon_e\right)}\right] = \\ &= \epsilon_e + \frac{1}{2}\left[\left(E_0 + \frac{2}{3}E_P\right) \pm \right. \\ &\left. \pm \sqrt{\left(E_0 + \frac{2}{3}E_P\right)^2 + \frac{8}{3}\epsilon E_P}\right] = \\ &= \epsilon_e + \frac{E_0 + \frac{2}{3}E_P}{2}\left[1 - \sqrt{1 + \frac{\frac{8}{3}\epsilon_e E_P}{\left(E_0 + \frac{2}{3}E_P\right)^2}}\right] \end{aligned} \tag{S151}$$

For sufficiently small energies ϵ, we can expand the square root. The derivation works exactly the same way as for small spin-orbit splitting, except that E_P needs to be replaced by $\frac{2}{3}E_P$. We obtain

$$\alpha_e = \frac{\left(\frac{2}{3}E_P\right)^2}{E_0\left(E_0 + \frac{2}{3}E_P\right)^2} \approx \frac{1}{E_0}, \tag{S152}$$

$$\alpha_e' = \frac{\left(\frac{2}{3}E_P\right)^2}{\left(E_0 + \frac{2}{3}E_P\right)^2} \approx 1, \tag{S153}$$

$$\beta_e = -\frac{2\left(\frac{2}{3}E_P\right)^3}{E_0\left(E_0 + \frac{2}{3}E_P\right)^4} \approx -\frac{3}{E_0 E_P}, \tag{S154}$$

$$\beta_e' = -\frac{2E_0\left(\frac{2}{3}E_P\right)^3}{\left(E_0 + \frac{2}{3}E_P\right)^4} \approx -\frac{3E_0}{E_P}. \tag{S155}$$

We find that the first-order nonparabolicity coefficient α_e' is still about one, but the second-order coefficient β_e' has increased by a factor of $\frac{3}{2}$. We expect the same convergence criteria $\epsilon \ll E_0$ and $E_0 \ll E_P$ as for small spin-orbit splittings.

We can use the same approach for the light hole band and find for large spin-orbit splittings

$$\alpha_{lh} = \frac{\left(\frac{2}{3}E_P\right)^2}{E_0\left(\frac{2}{3}E_P - E_0\right)^2} \approx \frac{1}{E_0}, \tag{S156}$$

$$\alpha_{lh}' = \frac{\left(\frac{2}{3}E_P\right)^2}{\left(\frac{2}{3}E_P - E_0\right)^2} \approx 1, \tag{S157}$$

$$\beta_{lh} = \frac{2\left(\frac{2}{3}E_P\right)^3}{E_0\left(\frac{2}{3}E_P - E_0\right)^4} \approx \frac{3}{E_0 E_P}, \tag{S158}$$

$$\beta_{lh}' = \frac{2E_0\left(\frac{2}{3}E_P\right)^3}{\left(\frac{2}{3}E_P - E_0\right)^4} \approx \frac{3E_0}{E_P}. \tag{S159}$$

These equations agree with Bartoli (1983), Eqs. (A3a-A6), in the limit of large spin-orbit splittings. For InSb, this expansion should produce better results than for GaAs, because the interaction of the light and split-off hole bands is weaker in InSb, see Fig. S6.

C. Using the electron/light hole reduced mass

For this section, it is important to use the masses obtained from the large spin-orbit splitting approximation consistently. If we mix and match the masses from different approximations, then Eq. (S163) will not be valid.

We can also start the calculation of the nonparabolicity parameters α and β with Eqs. (S44) and (S69) for the conduction band, which read

$$\epsilon_e = K^2 + \frac{E_0}{2}\left(\sqrt{1 + K^2\frac{2}{\mu_{lh}E_0}} - 1\right), \tag{S160}$$

where $\epsilon_e = E - E_0$ is again the energy of the electron above the conduction band minimum. This leads us to

the biquadratic equation

$$K^4 - K^2\left(E_0 + \frac{E_0}{2\mu_{lh}} + 2\epsilon_e\right) + \epsilon_e\left(\epsilon_e + E_0\right) = 0, \quad \text{(S161)}$$

which has the solution

$$K^2 = \epsilon_e + \frac{E_0}{2}\left(1 + \frac{1}{2\mu_{lh}}\right)\left[1 - \sqrt{1 + \frac{2\epsilon_e}{\mu_{lh}E_0\left(1 + \frac{1}{2\mu_{lh}}\right)^2}}\right]. \quad \text{(S162)}$$

We recognize the term in round parentheses as the inverse effective electron mass given by Eqs. (S51) and (S77). This simplies the expression for K^2 to read

$$K^2 = \epsilon_e + \frac{E_0}{2m_e^*}\left(1 - \sqrt{1 + \frac{2\epsilon_e m_e^{*2}}{\mu_{lh}E_0}}\right). \quad \text{(S163)}$$

By expanding the square root in orders of ϵ_e, we obtain

$$K^2 = \epsilon_e\left(1 - \frac{m_e^*}{2\mu_{lh}} + \frac{m_e^{*3}\epsilon_e}{4\mu_{lh}^2E_0} - \frac{m_e^{*5}\epsilon_e^2}{4\mu_{lh}^3E_0^2}\right). \quad \text{(S164)}$$

Equations (S51) and (S77) imply that the first two terms inside the parentheses are equal to the effective electron mass:

$$m_e^* = 1 - \frac{m_e^*}{2\mu_{lh}}. \quad \text{(S165)}$$

The wave vector as a function of energy can therefore be obtained from

$$K^2 = m_e^*\epsilon_e\left(1 + \frac{m_e^{*2}\epsilon_e}{4\mu_{lh}^2E_0} - \frac{m_e^{*4}\epsilon_e^2}{4\mu_{lh}^3E_0^2}\right). \quad \text{(S166)}$$

The nonparabolicity parameters for the electron band are therefore given by, with reference to Eq. (S165),

$$\alpha_e = \frac{m_e^{*2}}{4\mu_{lh}^2E_0} = \frac{\left(1 - m_e\right)^2}{E_0}, \quad \text{(S167)}$$

$$\alpha_e' = \frac{m_e^{*2}}{4\mu_{lh}^2} = \left(\frac{m_e^*}{2\mu_{lh}}\right)^2 = \left(1 - m_e^*\right)^2, \quad \text{(S168)}$$

$$\beta_e = -\frac{m_e^{*4}}{4\mu_{lh}^3E_0^2} = \frac{-2m_e^*\left(1 - m_e^*\right)^3}{E_0^2}, \quad \text{(S169)}$$

$$\beta_e' = -\frac{m_e^{*4}}{4\mu_{lh}^3} = -2m_e^*\left(1 - m_e^*\right)^3. \quad \text{(S170)}$$

These nonparabolicity parameters were derived for both the small and large spin-orbit splitting approximations. It is tempting to use them also for the case of general spin-orbit splittings, since they offer an elegant way of interpolating between both limits. We note that Eqs. (S168) and (S170) are the leading factors in the nonparabolicity expressions given by Bartoli (1983) in the limit of large spin-orbit splittings. Compare also Eqs. (1f) and (6a) in Palik, Picus, Teitler, and Wallis (1961). α_e' is approximately equal to unity, since the effective electron mass is small. β_e' is on the order of the effective electron mass, i.e., much smaller than α_e'.

For the light hole band, the energy is

$$E_{lh} = K^2 - \frac{E_0}{2}\left(\sqrt{1 + K^2\frac{2}{\mu_{lh}E_0}} - 1\right), \quad \text{(S171)}$$

which leads to the biquadratic equation

$$K^4 - K^2\left(-E_0 + \frac{E_0}{2\mu_{lh}} + 2E_{lh}\right) + E_{lh}\left(E_{lh} - E_0\right) = 0. \quad \text{(S172)}$$

This equation is similar to Eq. (S161) for the conduction band, except that both E_0 and μ_{lh} appear with the opposite sign. The solution is therefore

$$K^2 = E_{lh} - \frac{E_0}{2}\left(1 - \frac{1}{2\mu_{lh}}\right)\left[1 - \sqrt{1 + \frac{2E_{lh}}{\mu_{lh}E_0\left(1 - \frac{1}{2\mu_{lh}}\right)^2}}\right]. \quad \text{(S173)}$$

The term in round parentheses is related to the inverse light hole effective mass through Eqs. (S48) and (S74). This simplies the expression for K^2 to read

$$K^2 = E_{lh} + \frac{E_0}{2m_{lh}^*}\left[1 - \sqrt{1 + \frac{2E_{lh}m_{lh}^{*2}}{\mu_{lh}E_0}}\right]. \quad \text{(S174)}$$

By expanding the square root in orders of E_{lh}, we obtain

$$K^2 = E_{lh}\left(1 - \frac{m_{lh}^*}{2\mu_{lh}} + \frac{m_{lh}^{*3}E_{lh}}{4\mu_{lh}^2E_0} - \frac{m_{lh}^{*5}E_{lh}^2}{4\mu_{lh}^3E_0^2}\right). \quad \text{(S175)}$$

Equations (S48) and (S74) imply that the first two terms inside the parentheses are related to the effective light hole mass by

$$-m_{lh}^* = 1 - \frac{m_{lh}^*}{2\mu_{lh}}. \quad \text{(S176)}$$

If we remember the definition $\epsilon_{lh} = -E_{lh}$ from the beginning of Sec. S11, then the wave vector as a function of energy can be obtained from

$$K^2 = m_{lh}^*\epsilon_{lh}\left(1 + \frac{m_{lh}^{*2}\epsilon_{lh}}{4\mu_{lh}^2E_0} + \frac{m_{lh}^{*4}\epsilon_{lh}^2}{4\mu_{lh}^3E_0^2}\right). \quad \text{(S177)}$$

Equations (S166) and (S177) are not identical, if we change the band subscripts for the energy and for the effective mass (but not for the reduced mass), because they have the opposite sign for the quartic nonparabolicity parameter β'.

The nonparabolicity parameters for the light hole band are therefore given by, with reference to Eq. (S176),

$$\alpha_{lh} = \frac{m_{lh}^{*2}}{4\mu_{lh}^2E_0} = \frac{\left(1 + m_{lh}\right)^2}{E_0}, \quad \text{(S178)}$$

$$\alpha_{lh}' = \frac{m_{lh}^{*2}}{4\mu_{lh}^2} = \left(\frac{m_{lh}^*}{2\mu_{lh}}\right)^2 = \left(1 + m_{lh}^*\right)^2, \quad \text{(S179)}$$

TABLE SI. Effective masses and nonparabolicity parameters for InSb at low temperature, calculated with the values E_0=0.237 eV, Δ_0=0.81 eV, and E_P=23.1 eV taken from Lawaetz (1971). The top row uses the limit of large spin-orbit splittings in Sec. S7 with Eqs. (S65,S66,S68,S168,S170,S179,S181). The second row uses the perturbative approximation in Sec. S9 with Eqs. (S110,S118,S119,S168,S170,S179,S181). The third row lists results with finite spin-orbit splittings from the Appendix in Bartoli (1983).

m_e^*	m_{lh}^*	μ_{lh}	α_e'	β_e'	α_{lh}'	β_{lh}'
0.0152	0.0156	0.00769	0.970	−0.029	1.032	0.033
0.0136	0.0156	0.00728	0.973	−0.026	1.032	0.033
0.0136	0.0156	0.00728	0.896	−0.075	1.182	0.264

$$\beta_{lh} = \frac{m_{lh}^{*4}}{4\mu_{lh}^3 E_0^2} = \frac{2m_{lh}^* \left(1+m_{lh}^*\right)^3}{E_0^2}, \tag{S180}$$

$$\beta_{lh}' = \frac{m_{lh}^{*4}}{4\mu_{lh}^3} = 2m_{lh}^* \left(1+m_{lh}^*\right)^3. \tag{S181}$$

Equations. (S179) and (S181) are the leading factors in the nonparabolicity expressions given by Bartoli (1983) in the limit of large spin-orbit splittings.

The expressions given in this section are evaluated as shown in Fig. S11 for the electron and light hole bands of InSb. The parabolic approximation (dotted) is poor except for very small energies above the conduction band minimum or below the valence band maximum. Keeping only the lowest-order nonparabolic term α_i is a good approximation to the exact expression containing the square root. If the β_i-term is included (dot-dashed), the deviation from the exact square-root expression (S163) is nearly indistinguishable on this scale. This energy dispersion is quite similar to that calculated by Kane (1957), Fig. 1, using an 8×8 $\vec{k}\cdot\vec{p}$-model with corrections due to interactions with higher-energy bands.

While the light hole mass of m_{lh}=0.0156 given by the large spin-orbit splitting approximation is similar to the experimental mass, the electron mass of m_e=0.0152 is about 10% too high. But inserting the experimental electron mass (m_e=0.013 or 0.014) in Eq. (S163) yields inconsistent results, because different approximations are mixed in the expressions, with invalid outcomes.

Table SI shows the effective electron and light hole masses and nonparabolicity parameters calculated in different approximations.

D. Bartoli's expressions

Without much discussion and no derivation or references, Bartoli *et al.* (1983) give the following expressions for the effective masses and the non-parabolicity:

$$\delta = \frac{\Delta_0}{E_0}, \tag{S182}$$

$$\frac{1}{m_e^*} = 1 + \frac{E_P}{3}\left(\frac{2}{E_0} + \frac{1}{E_0+\Delta_0}\right), \tag{S183}$$

$$m_l^* = \frac{(\delta+1.5)\, m_e^*}{\delta + 1 - m_e^*\,(2\delta+2.5)}, \tag{S184}$$

$$\alpha_e = \frac{\left(1-m_e^*\right)^2\left(1.5+2\delta+\delta^2\right)}{(1.5+\delta)\,(1+\delta)}, \tag{S185}$$

$$\beta_e = -\frac{2\left(1-m_e^*\right)^3}{(1.5+\delta)^2\,(1+\delta)}, \tag{S186}$$

$$\alpha_{lh} = \left(1+m_{lh}^*\right)^2\left(1+\frac{1}{2\delta}\right), \tag{S187}$$

$$\beta_{lh} = \left(1+m_{lh}^*\right)^3\left[\frac{1}{2\delta} + \frac{3}{4\delta^2} + 2m_{lh}^*\left(1+\frac{1}{2\delta}\right)^2\right] \tag{S188}$$

They mention that Eq. (S185) was previously given by Palik *et al.* (1961).

The square of the wave vector as function of the excess energy ϵ_i evaluated using the Bartoli expressions is shown in Fig. S12. The splitting between the electron and light hole bands is larger for these expressions than for the large spin-orbit splitting approximation, because the second-order corrections involving the β_i-terms are larger. Overall, the agreement with the bands calculated by Kane (1957) is worse than for the pure approximation for large spin-orbit splittings. Therefore, it does not seem beneficial to use the Bartoli expressions for InSb.

E. Expression given by Menendez (2020)

Menendez (2020) introduced a nonparabolicity parameter for the conduction band

$$\Delta_\Gamma = \frac{3}{2}\left(\frac{2}{E_0} + \frac{1}{E_0+\Delta_0}\right)^{-1}. \tag{S189}$$

In our notation, $2\Delta_\Gamma = 1/\alpha_e$. Therefore, Δ_Γ should approach $E_0/2$ for both small and large spin-orbit splittings. This condition is fulfilled for small spin-orbit splittings, but not for large spin-orbit splittings, where $\Delta_\Gamma = 3E_0/4$. Therefore, the use of Δ_Γ for InSb is not recommended, because it will underestimate the nonparabolicity by up to 30%.

S12. DENSITY OF STATES

For the calculation of the chemical potential, we need the density of states

$$g_i\left(\epsilon_i\right) = \frac{1}{4\pi^3}\int d^3\vec{k}\,\delta\left(E_{i\vec{k}} - \epsilon_i\right) = \frac{1}{\pi^2}\int_0^\infty k^2 dk\,\delta\left(E_{ik} - \epsilon_i\right), \tag{S190}$$

where i is the band index. We have included the spin degeneracy and assume that the bands are spherically symmetric.

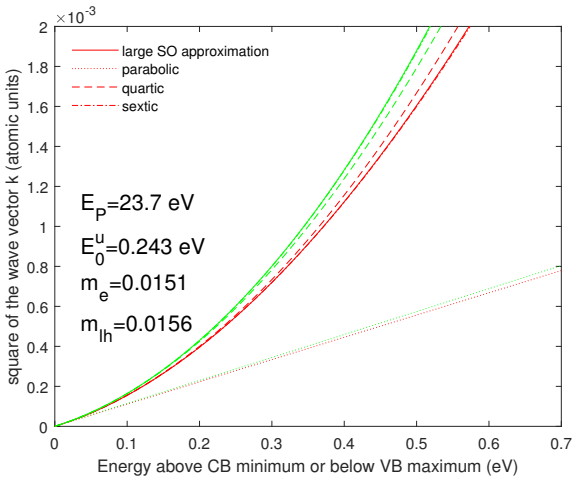

FIG. S11. The square of the wave vector (in atomic units) versus energy from the extremum for electrons (red) and light holes (green) in InSb, calculated using Eq. (S163) within the large spin-orbit splitting approximation (solid), using masses from Table SI. The dotted lines show the parabolic band dispersion. The dashed and dot-dashed lines show the next two terms in the Taylor expansions of the square root. Only the first term (β_i=0) gives a good approximation (dashed). If the β_i-term is included (dot-dashed), the deviation from the exact square-root expression (S163) is nearly indistinguishable on this scale.

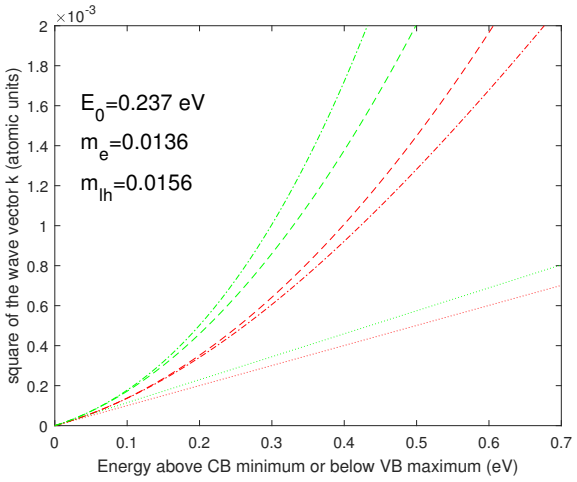

FIG. S12. As Fig. S11, but evaluated using the expressions given by Bartoli *et al.* (1983).

By taking the derivative of Eq. (S127) on both sides, we find

$$dk = \sqrt{\frac{m_0 m_i}{2\hbar^2 \epsilon_i}} \frac{1 + 2\alpha_i \epsilon_i + 3\beta_i \epsilon_i^2}{\sqrt{1 + \alpha_i \epsilon_i + \beta_i \epsilon_i^2}} d\epsilon_i, \tag{S191}$$

compare Bartoli (A7), and

$$k^2 dk = \frac{1}{2} \left(\frac{2 m_0 m_i}{\hbar^2} \right)^{\frac{3}{2}} \times$$
$$\times \sqrt{\epsilon_i \left(1 + \alpha_i \epsilon_i + \beta_i \epsilon_i^2\right)} \left(1 + 2\alpha_i \epsilon_i + 3\beta_i \epsilon_i^2\right) d\epsilon_i. \tag{S192}$$

The density of states is therefore

$$g_i(\epsilon_i) = \frac{1}{2\pi^2} \left(\frac{2 m_0 m_i}{\hbar^2} \right)^{\frac{3}{2}} \times$$
$$\times \sqrt{\epsilon_i \left(1 + \alpha_i \epsilon_i + \beta_i \epsilon_i^2\right)} \left(1 + 2\alpha_i \epsilon_i + 3\beta_i \epsilon_i^2\right), \tag{S193}$$

compare Fox (3.24) and Masut (2022), Eq. (6). We see that the nonparabolicity enhances the density of states by a factor

$$\left(1 + 2\alpha_i \epsilon_i + 3\beta_i \epsilon_i^2\right) \sqrt{1 + \alpha_i \epsilon_i + \beta_i \epsilon^2} \approx 1 + \frac{5}{2} \alpha_i \epsilon_i \tag{S194}$$

to first order in $\alpha_i \epsilon_i$ if we set β_i to zero. Since the density of states depends on $m_i^{\frac{3}{2}}$, we can define an energy-dependent density-of-states effective mass

$$m_{i,\mathrm{DOS}}(\epsilon_i) = m_i \sqrt[3]{1 + \alpha_i \epsilon_i + \beta_i \epsilon_i^2} \left(1 + 2\alpha_i \epsilon_i + 3\beta_i \epsilon_i^2\right)^{\frac{2}{3}}. \tag{S195}$$

By setting β_e=0 and keeping only terms linear in $\alpha_e \epsilon_e$, the effective electron mass m_e increases approximately like

$$m_{e,\mathrm{DOS}}(\epsilon) \approx m_e \left(1 + \frac{5}{3} \alpha_e \epsilon_e\right). \tag{S196}$$

In other words, when the excess energy ϵ_e is equal to the band gap E_0 (i.e., $\alpha_e \epsilon_e \approx 1$), the effective electron mass m_e has nearly tripled. This is shown in Fig. S13, which plots the effective density-of-states electron mass of InSb as a function of excess energy above the conduction band minimum. Most of the mass enhancement is due to α_e-term (shown by the dotted line). We therefore have confidence that the expansion (S193) converges well in the approximation for large spin-orbit splittings. The energy dependence of the effective density-of-states light hole mass is also shown in Fig. S13.

S13. CHEMICAL POTENTIAL WITH NONPARABOLICITY

A. General expressions for all semiconductors

We apply the density of states (S193) for nonparabolic bands to calculate the chemical potential μ and the intrinsic carrier concentration n for InSb as a function of temperature T.

The electron density n at temperature T is given by

$$n(T) = \int_0^\infty d\epsilon g_e(\epsilon) f(E_0 + \epsilon_e), \tag{S197}$$

where $f(E)$ is the Fermi-Dirac distribution function

$$f(E) = \frac{1}{\exp\left(\frac{E-\mu}{k_B T}\right) + 1}. \tag{S198}$$

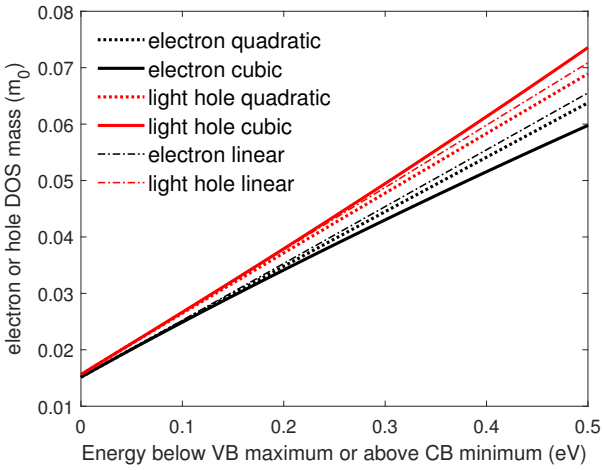

FIG. S13. Effective density of states electron mass m_e (black) and light hole mass m_{lh} (red) of InSb at 0 K as a function of excess energy above the conduction band minimum, calculated from Eq. (S195) (solid). The dotted lines show the results with β_e or β_{lh} set to zero. The dash-dotted lines show the linear expansion (S196).

By setting β_e=0 and keeping only terms linear in α_e, we find that the density of states enhancement factor (S194) is approximately $1+\frac{5}{2}\alpha_e\epsilon_e$. With the substitutions $y=\epsilon_e/k_BT$ and $x=(\mu-E_0)/k_BT$, the electron density can be written using Fermi-Dirac integrals as (compare Menendez 2020)

$$\begin{aligned} n(T) &= N_e(T)\left[F_{\frac{1}{2}}\left(\frac{\mu-E_0}{k_BT}\right)+\right. \\ &+ \left.\frac{15}{4}\alpha_e k_BT F_{\frac{3}{2}}\left(\frac{\mu-E_0}{k_BT}\right)\right], \end{aligned} \tag{S199}$$

where the prefactor N_e is given by

$$N_e(T)=\frac{1}{4}\left(\frac{2m_0m_ek_BT}{\pi\hbar^2}\right)^{3/2}. \tag{S200}$$

Similarly, the light and heavy hole densities are given by

$$\begin{aligned} p_{lh}(T) &= N_{lh}(T)\left[F_{\frac{1}{2}}\left(-\frac{\mu}{k_BT}\right)+\right. \\ &+ \left.\frac{15}{4}\alpha_{lh}k_BTF_{\frac{3}{2}}\left(-\frac{\mu}{k_BT}\right)\right], \\ p_{hh}(T) &= N_{hh}(T)F_{\frac{1}{2}}\left(-\frac{\mu}{k_BT}\right), \end{aligned} \tag{S201}$$

if we ignore the nonparabolicity of the heavy hole band, with prefactors given by

$$\begin{aligned} N_{lh}(T) &= \frac{1}{4}\left(\frac{2m_0m_{lh}k_BT}{\pi\hbar^2}\right)^{3/2} \quad \text{and} \\ N_{hh}(T) &= \frac{1}{4}\left(\frac{2m_0m_{hh}k_BT}{\pi\hbar^2}\right)^{3/2}. \end{aligned} \tag{S202}$$

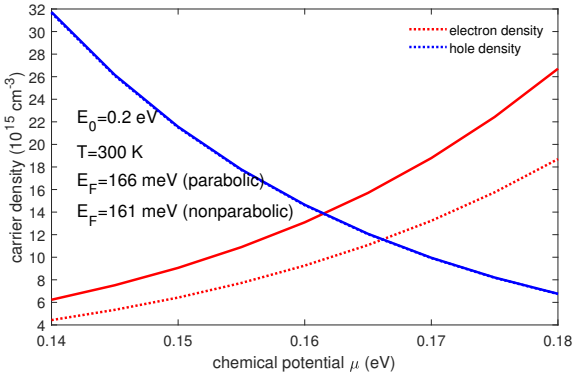

FIG. S14. Electron (red) and hole density (blue) as a function of chemical potential at 300 K in the parabolic approximation (dotted) and with the lowest nonparabolic corrections (solid). The band gap and heavy hole mass were assumed to be 0.2 eV and 0.43, respectively, and the electron and light hole mass were calculated in the large spin-orbit approximation.

We can find the chemical potential by solving the equation

$$n(T)=p_{lh}(T)+p_{hh}(T) \tag{S203}$$

at a given temperature, for example using polylogarithm functions in MATLAB. As an example, we show the electron and hole density of InSb at 300 K as a function of chemical potential in Fig. S14. For the holes, the light hole density is only a very small contribution, because the heavy hole is about 30 times heavier than the light hole. Therefore, the nonparabolicity correction for the hole bands does not matter much. The nonparabolicity correction for the electron concentration is sizeable, which can be seen from the difference between the red dotted and solid lines.

B. Evaluation of several scenarios for InSb

At 300 K with a band gap of 0.2 eV, E_P=23.1 eV, a heavy hole mass of 0.43, and electron and light hole masses calculated within the large spin-orbit splitting approximation, electron and hole concentrations become equal at a chemical potential of 161 meV (166 meV) with (without) the lowest order nonparabolic correction and α_e and α_{lh} from Table SI. Compare Rivero Arias (2023).

Two trends are seen in Fig. S14: First, the chemical potential is lower when the nonparabolicity is taken into account. Second, the carrier density (at this lower chemical potential) is higher than in the parabolic approximation. Both trends can be understood from the nondegenerate (classical Maxwell-Boltzmann) limit for parabolic bands (Sze 1981, Ashcroft & Mermin 1976)

$$\begin{aligned} \mu &= \frac{E_0}{2}+\frac{3kT}{4}\ln\left(\frac{m_{\rm dv}}{m_e}\right), \\ n &= p= \end{aligned} \tag{S204}$$

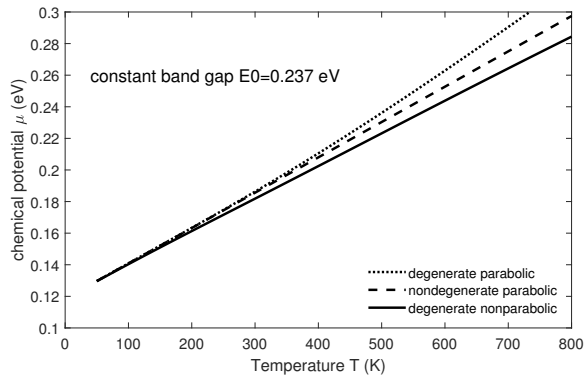

FIG. S15. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. A band gap of 0.237 eV and temperature-independent effective masses were assumed.

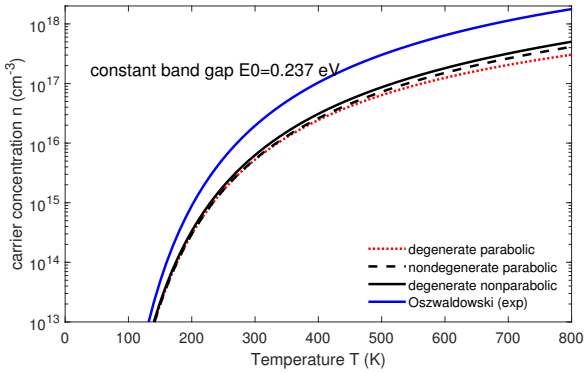

FIG. S16. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. A band gap of 0.237 eV and temperature-independent effective masses were assumed.

$$= 2\left(\frac{m_0 k_B T}{2\pi\hbar^2}\right)^{3/2} (m_e m_{dv})^{3/4} \exp\left(-\frac{E_0}{2k_B T}\right) \quad \text{(S205)}$$

If the electron mass increases due to nonparabolicity while the density-of-states hole mass m_{dv} remains nearly constant, then the chemical potential decreases following Eq. (S204). At the same time, the carrier concentration increases like $m_e^{3/4}$ due to Eq. (S205).

Figures S15 and S16 show the chemical potential and the intrinsic carrier concentration versus temperature, assuming a band gap and effective masses that are independent of temperature. We again see that the chemical potential is lowered and the carrier concentration increased by nonparabolicity.

In our next model for the temperature dependence of the chemical potential and the intrinsic carrier concentration, we also include the full temperature dependence of the band gap (electron-phonon and thermal expansion contributions) using the Bose-Einstein expression

$$E_0(T) = E_B - a_B\left[1 + \frac{2}{\exp(\Omega/k_B T) - 1}\right] \quad \text{(S206)}$$

with parameters E_B=261 meV, a_B=26 meV, and Ω=18.9 meV (correcting an error in the article by Rivero Arias, 2023). We ignore the small temperature dependence of the coupling parameter E_P due to thermal expansion, see Eq. (21) in Emminger (2022). The lower band gap at high temperatures will also reduce the light hole and electron masses according to Eq. (S31) and (S110). The temperature dependence of the light hole mass is not important, because the hole density is dominated by heavy holes. The effective mass of heavy holes should not change much with temperature, because this mass is dominated by the coupling with higher-energy conduction bands (see Dresselhaus, Kip, and Kittel, for example). The higher gaps only change by a small fraction. The temperature dependence of the heavy hole mass (both theory and experiment) for InSb has been discussed by Oszwaldowski and Zimpel (1988). For our purposes, we select a constant value of the heavy hole mass (independent of temperature) and parabolic heavy hole bands.

The choice of the effective masses has been discussed in the literature. Some authors, for example Oszwaldowski and Zimpel (1988), have argued that the unrenormalized band gap E_B should be used to calculate the effective masses within $\vec{k}\cdot\vec{p}$-theory. At the moment, however, we take the view that the experimental band gap E_0 (sometimes also called the thermal gap or optical gap) should influence the effective light hole and electron masses, with a nearly constant matrix element E_P. The effective masses therefore decrease significantly with increasing temperature as the band gap closes according to Eq. (S206).

As shown in Fig. S17, the chemical potential increases superlinearly at high temperatures, where the band gap shrinks significantly, see Eq. (S206) and Fig. 3 in Rivero Arias (2023). The band gap reduction also decreases the effective electron mass, see Eq. (S31), and therefore increases the chemical potential according to Eq. (S204). The temperature dependence of the intrinsic carrier concentration is shown in Fig. S18 with the temperature dependence of the band gap and the effective masses considered. At 800 K, the carrier concentration is increased by about 33% compared to the calculation with a temperature-independent band gap and electron mass shown in Fig. S16, from 6×10^{17} to 8×10^{17} cm^{-3}.

Oszwaldowski and Zimpel (1988) obtained the temperature dependence of the intrinsic carrier concentration of InSb from 200 to 800 K from Hall measurements, also shown in Fig. S18. Assuming a Hall factor of unity, they found an intrinsic carrier concentration near 1.8×10^{18} cm^{-3} at 800 K, about two times higher than the carrier concentration of 8.2×10^{17} cm^{-3} cal-

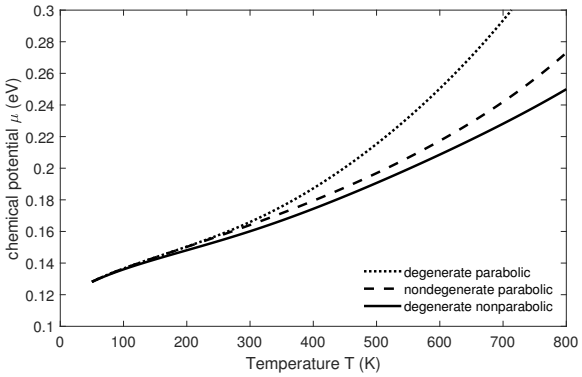

FIG. S17. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) and of the light hole and electron masses was included.

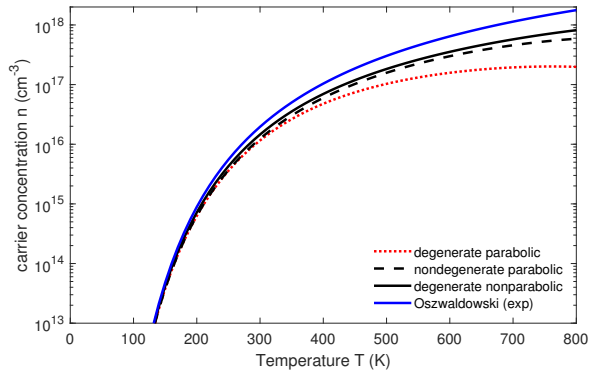

FIG. S18. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) and of the light hole and electron masses was included. The blue line shows a fit to carrier concentrations determined from Hall measurements by Oszwaldowski and Zimpel (1988).

culated with our degenerate nonparabolic model (with temperature-dependent band gap and effective masses). Only a much higher (perhaps temperature-independent) density of states mass for the electrons would explain these experimental data. The next order nonparabolicity correction would not improve the agreement, because $\beta_e<0$.

For completeness, we show the temperature dependence of the band gap and of the effective electron and light hole masses in Fig. S19. As expected, all have a similar temperature dependence. It is interesting that the chemical potential at 800 K is about five times as large as the band gap.

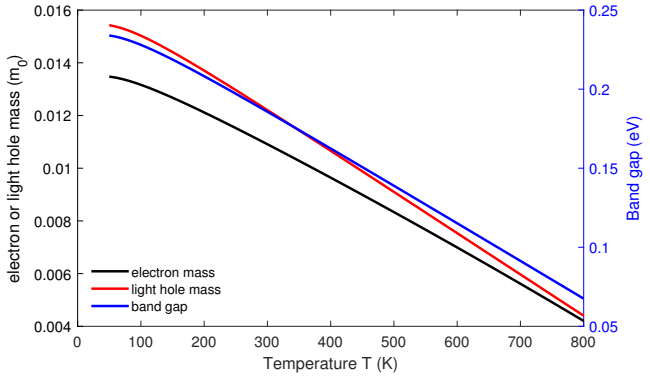

FIG. S19. Effective masses of the electron (black) and light hole (red) bands and band gap (blue) of InSb as a function of temperature.

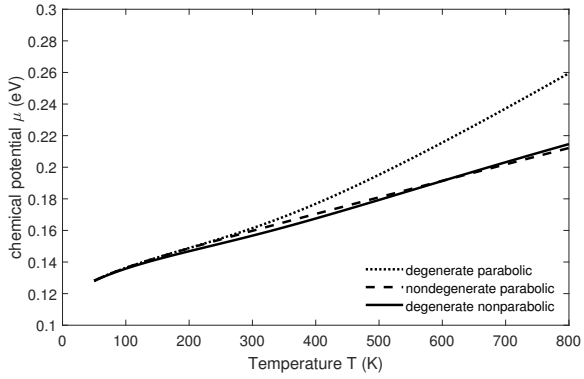

FIG. S20. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses and nonparabolicity parameters were kept constant at values calculated from the low-temperature band gap.

We also calculated the temperature dependence of the chemical potential and the intrinsic carrier concentration of InSb by assuming that the thermal or optical gap (which enters the Fermi-Dirac distribution) varies with temperature, while the effective masses and nonparabolicity parameters are calculated using $\vec{k}\cdot\vec{p}$-theory from the low-temperature band gap. The results are shown in Figs. S20 and S21. With this model, the calculated carrier density at 800 K is only 20% lower than the experimental value measured by Oszwaldowski and Zimpel (1988). It is also a remarkable coincidence that the degenerate nonparabolic carrier concentration is nearly the same as the nondegenerate parabolic result.

Next, we will calculate the chemical potential and the intrinsic carrier concentration using the thermal (or optical) gap from Eq. (S206) in the Fermi-Dirac distribution

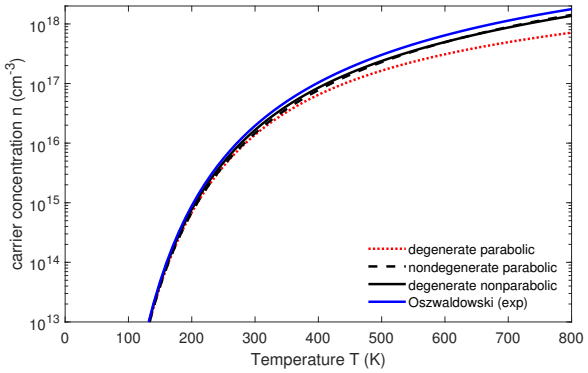

FIG. S21. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses and nonparabolicity parameters were kept constant at values calculated from the low-temperature band gap. The blue line shows a fit to carrier concentrations determined from Hall measurements by Oszwaldowski and Zimpel (1988).

function, while considering only the thermal expansion contribution to the band gap changes in the calculation of the effective masses and the nonparabolicity parameters, but not the redshift due to electron-phonon renormalization. The redshift of the direct gap due to thermal expansion is given by (Zollner 1991 GaSb)

$$\left(\frac{\partial E_0}{\partial T}\right)_{\mathrm{TE}} = -3\alpha\left(T\right) B \left(\frac{\partial E_0}{\partial p}\right)_T, \tag{S207}$$

where $\alpha\left(T\right)$ is the temperature-dependent thermal expansion coefficient taken from S. I. Novikova, *Thermal expansion*, in *Semiconductors and Semimetals*, vol. 2, edited by R. K. Willardson and A. C. Beer, pp. 33-38 (1966), B=46 GPa the bulk modulus (assumed to be independent of temperature, LaBoe Vol. 17a, page 317), and the last factor equal to 0.155 eV/GPa is the pressure coefficient of the InSb band gap (Booth 1982, LaBoe Vol. 22a, page 124). For $\alpha\left(T\right)$ above room temperature, see Cai J. Appl. Phys. 2013.

The temperature dependence of the band gap due to thermal expansion and ignoring the electron-phonon renormalization is therefore

$$E_0^{\mathrm{TE}}\left(T\right) = E_0\left(T=0\right) - 3B\left(\frac{\partial E_0}{\partial p}\right)_T \int_0^T \alpha\left(\theta\right) d\theta. \tag{S208}$$

We also call this the "mass band gap".

Following Roucka (PRB 81, 245214, 2010), we write the thermal expansion coefficient of zinc blende semiconductors as

$$\alpha\left(T\right) = A\left(\frac{T}{\Theta_D}\right)^3 I_D\left(\frac{\Theta_D}{T}\right), \tag{S209}$$

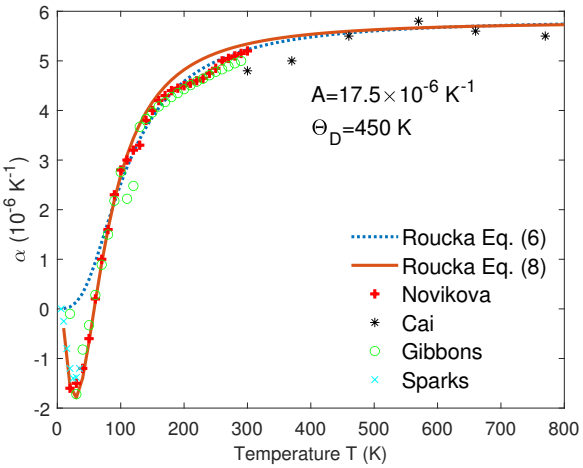

FIG. S22. Linear thermal expansion coefficient α versus temperature taken from Novikova (1966) (symbols) along with a fit to the data using Eq. (S209) (solid).

where

$$I_D\left(x_D\right) = \int_0^{x_D} \frac{x^4 e^x dx}{\left(e^x - 1\right)^2} \tag{S210}$$

is the Debye integral (which can be solved numerically in MATLAB), A is an adjustable parameter, and Θ_D the Debye temperature. The Debye temperature for InSb is about 205 K for InSb (LaBo 22a 1987, page 129), but we treat it as an adjustable parameter to fit the thermal expansion coefficients from Novikova (1966). With parameters A=17.5×10^{-6} K^{-1} and Θ_D=450 K, satisfactory agreement with the experimental data can be achieved, see Fig. S22.

The temperature dependence of the band gap due to thermal expansion, calculated from Eq. (S208), is shown in Fig. S23. This is the same result as shown by Zollner (SSC, 1991).

The chemical potential and intrinsic carrier concentration of InSb as a function of temperature, taking into account only the thermal expansion contribution to the band gap shift (not the electron-phonon renormalization) when calculating the effective masses as a function of temperature, are shown in Figs. S24 and S25. Figure S26 shows the effective light hole and electron masses of InSb, the experimental band gap, and the thermal expansion band gap used to calculate the effective masses. As expected, the temperature dependence of the masses follows the dependence of the band gap due to thermal expansion in this situation. As shown by Fig. S25, the intrinsic carrier concentration is about 10^{18} cm^{-3}, lower than the result when keeping the masses constant as in Fig. S21 and higher than when calculating the temperature dependence of the masses from the experimental band gap as shown in Fig. S18.

Of course, this model for the temperature dependence of the intrinsic carrier concentration is not consistent,

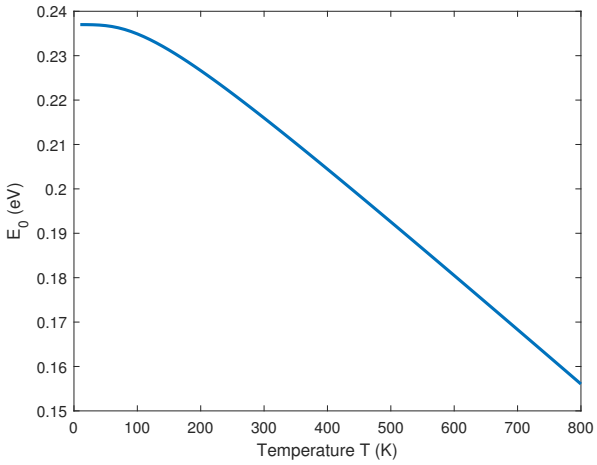

FIG. S23. Termperature dependence of the direct band gap of InSb due to thermal expansion. The electron-phonon contribution has been ignored. Compare Zollner (SSC, 1991).

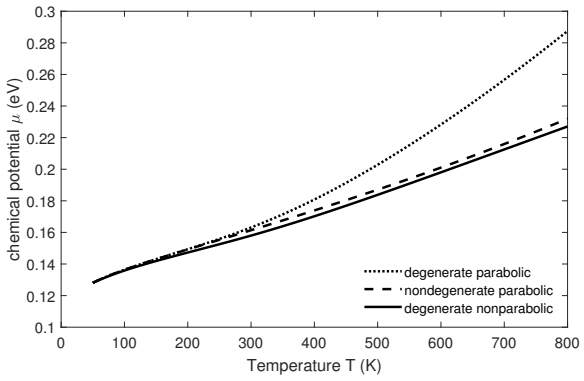

FIG. S24. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses were calculated taking into account only the thermal expansion contribution to the band gap, not the electron-phonon renormalization.

because the experimental low-temperature band gap, including the electron-phonon renormalization due to zero-point phonon oscillations, was used to calculate the low-temperature masses, but the electron-phonon renormalization was ignored in the temperature dependence of the effective masses. To build a consistent $\vec{k}\cdot\vec{p}$ model, we need to remove the zero-point electron phonon contribution to the low-temperature effective masses. We determine a new $\vec{k}\cdot\vec{p}$ momentum matrix parameter as follows:

Within Kane's 8×8 $\vec{k}\cdot\vec{p}$, the effective masses of the electron and light hole bands depend on three parameters, the band gap E_0, the spin-orbit splitting Δ_0, and

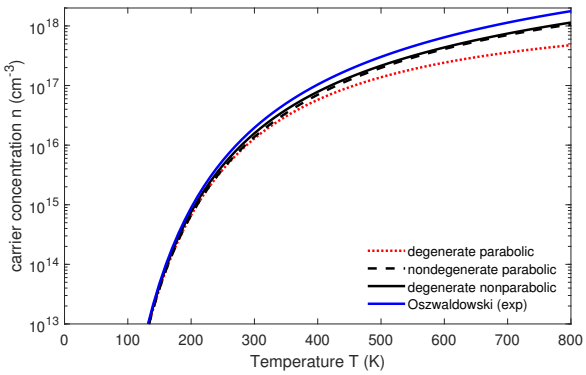

FIG. S25. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses were calculated taking into account only the thermal expansion contribution to the band gap, not the electron-phonon renormalization. The blue line shows a fit to carrier concentrations determined from Hall measurements by Oszwaldowski and Zimpel (1988).

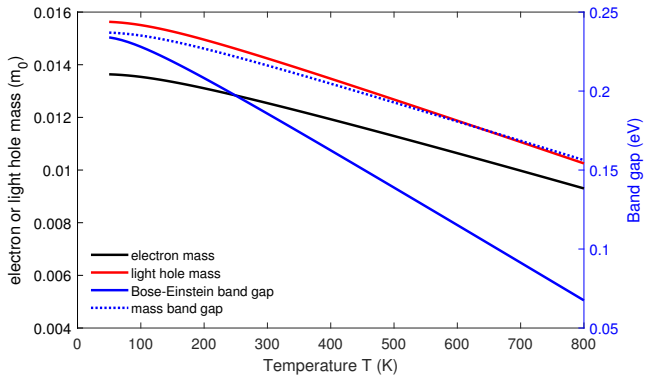

FIG. S26. Effective masses of the electron (black) and light hole (red) bands and band gap (blue) of InSb as a function of temperature, calculated taking into account only the thermal expansion contribution to the band gap, not the electron-phonon renormalization. The experimental gap and the thermal expansion contribution to the band gap shift are also shown.

the dipole matrix element parameter E_P through Eqs. (S31) and (S66). These expressions are accurate within the 8×8 model. Instead of the experimental band gap, we must use the unrenormalized band gap E_B=0.261 eV, see Eq. (S206), to calculate the effective masses. Starting with the light hole effective mass m^*_{lh}=0.0156 and solving Eq. (S66) for E_P, we obtain

$$E_P = \frac{3}{2}E_B\left(1+\frac{1}{m^*_{lh}}\right) = 25.5 \text{ eV}, \qquad \text{(S211)}$$

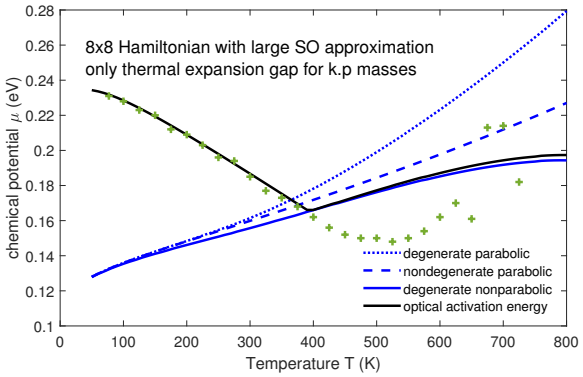

FIG. S27. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses were calculated taking into account only the thermal expansion contribution to the band gap, not the electron-phonon renormalization. The unrenormalized low-temperature gap was used to calculate the momentum matrix element from the experimental light-hole effective mass. The experimental direct band gap from a fit to the temperature-dependent infrared dielectric function with a Johs-Herzinger parametric oscillator model (symbols) are also shown, compare Rivero Arias (2023).

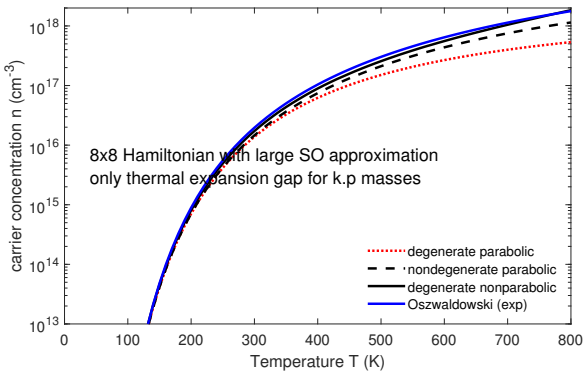

FIG. S28. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses were calculated taking into account only the thermal expansion contribution to the band gap, not the electron-phonon renormalization. The unrenormalized low-temperature gap was used to calculate the momentum matrix element from the experimental light-hole effective mass. The blue line shows a fit to carrier concentrations determined from Hall measurements by Oszwaldowski and Zimpel (1988).

somewhat larger than the usual value of 23.1 eV published by Lawaetz (1971). The corresponding effective electron mass with Δ_0=0.81 eV

$$\frac{1}{m_e^*} = 1 + \frac{E_P}{3}\left(\frac{2}{E_B} + \frac{1}{E_B + \Delta_0}\right) \qquad \text{(S212)}$$

equals 0.0135, which is in excellent agreement with the experimental value.

We therefore repeat the calculation of the temperature dependence of the chemical potential and the intrinsic carrier concentration using this revised value of E_P=25.5 eV. A similar approach was taken by V. G. Orlov and G. S. Sergeev (*Numerical simulation of the transport properties of indium antimonide*, Physics of the Solid State **55**, 2215-2222, 2013), but they do not describe how the use of the unrenormalized gap affects the value of the momentum matrix element. Masut (2022) also uses a similar approach.

The chemical potential and intrinsic carrier concentration of InSb as a function of temperature are shown in Figs. S27 and S28. Compare Fig. 1 in Masut (2022). The agreement between the experimental data of Oszwaldowski and Zimpel (1988) and our model is better than it should be. We need to remember that the large spin-orbit approximation overestimates the electron mass, because Δ_0 is taken to be infinite. Therefore, this model overestimates the intrinsic carrier concentration as well. There is some uncertainty in the temperature dependence of the

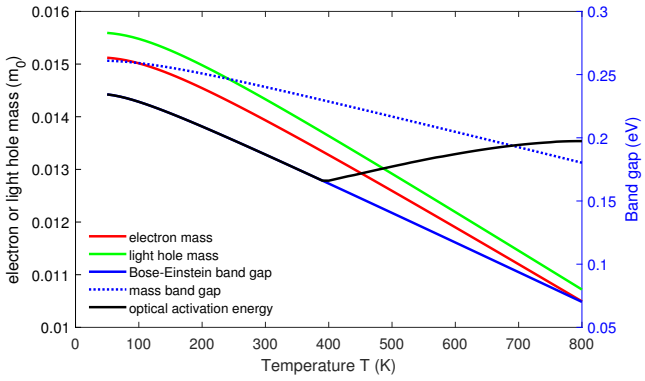

FIG. S29. Effective masses of the electron (black) and light hole (red) bands and band gap (blue) of InSb as a function of temperature, calculated taking into account only the thermal expansion contribution to the band gap, not the electron-phonon renormalization. The experimental gap and the thermal expansion contribution to the band gap shift are also shown.

direct gap, the thermal expansion coefficient, and the density of states heavy hole mass (which is taken to be spherical and parabolic). Therefore, perfect agreement between a model and the experimental data should not be expected.

All of the scenarios described above to calculate the temperature dependence of the effective masses have some flaws, especially how to treat the impact of the

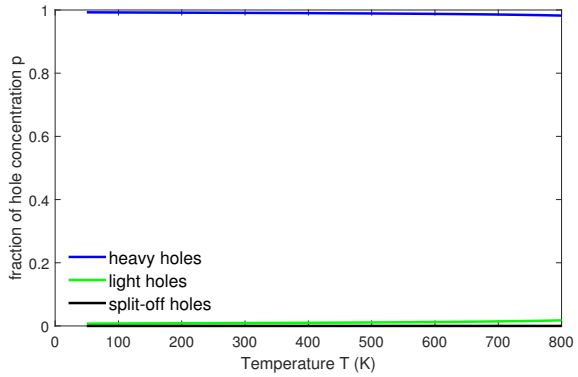

FIG. S30. Fraction of holes in the split-off, light hole, and heavy hole bands, calculated using the expressions in Sec. III B. At high temperatures, the effective light hole mass increases and therefore the population of the light hole band becomes noticable on this scale, but still very small.

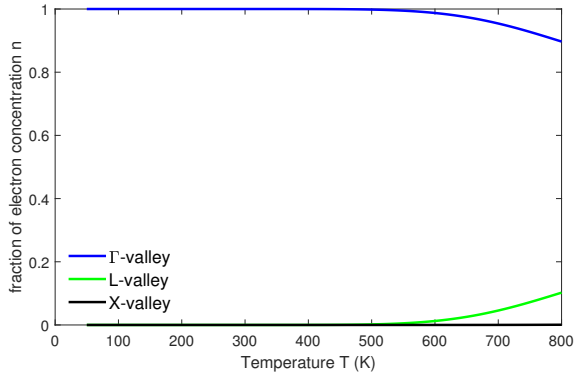

FIG. S31. Fraction of electrons in the Γ-, L-, and X-valleys of the conduction band, calculated using the expressions in Sec. III B. At low and medium temperatures, all electrons are in the Γ-valley, but up to 10% of electrons can be found in the L-valley at the highest temperatures. This was not considered in the analysis of the Hall experiments by Oszwaldowski and Zimpel.[23]

zero-point phonon oscillations on the renormalization of the band gap. The most consistent way to handle this matter is described in the main text of this review. The scenarios described in this section can be considered as evidence that treating the band gap as a constant (independent of temperature) or using the full experimental temperature dependence of the band gap $E_0^{\rm exp}$ to calculate the effective masses is not in agreement with Hall effect measurements.

The optical activation energy (i.e., the band gap observed in an optical absorption or ellipsometry experiment) is increased through the Burstein-Moss shift and given by

$$E_A = \max\left[E_0, E_0 + \left(1 + \frac{m_e}{m_h}\right)(\mu - E_0)\right], \quad \text{(S213)}$$

see Chakraborty, Datta, and Ghatak (Physica B **49**, 1179, 2003). The optical activation energy is equal to E_0 if the Fermi level is below the conduction band minimum, but increases as the Fermi level moves into the conduction band. The ratio of the masses takes into account that direct optical interband transitions are not possible at k=0, if the Fermi level is larger than the band gap. This optical activation energy needs to be compared to the ellipsometry data given by Rivero Arias *et al.*[38] as shown in the main text.

S14. SUPPLEMENTARY FIGURES FROM THE MAIN TEXT

The percentages of carriers in the various hole and electron bands as a function of temperature are shown in Figs. S30 and S31.