JVST A
Journal of Vacuum Science & Technology A

ARTICLE

pubs.aip.org/avs/jva

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

Cite as: J. Vac. Sci. Technol. A 43, 012801 (2025); doi: 10.1116/6.0003929
Submitted: 23 July 2024 · Accepted: 4 November 2024 ·
Published Online: 13 December 2024



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

AFFILIATIONS

[1]Department of Physics, New Mexico State University, MSC 3D, PO Box 30001, Las Cruces, New Mexico 88003-8001
[2]Department of Physics, Arizona State University, Tempe, Arizona 85287-1504

Note: This paper is part of the Special Topic Collection Commemorating the Career of Gerry Lucovsky.
[a]Author to whom correspondence should be addressed: zollner@nmsu.edu; URL: http://femto.nmsu.edu

ABSTRACT

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 k · 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 bandgap must be utilized. This includes only the redshift of the bandgap due to thermal expansion but 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 bandgaps 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 bandgap (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 bandgap $E_0^u$. This is not necessarily true in semiconductors with small bandgaps, such as InSb, for which higher-order terms must also be considered. That is the topic of this review.

Remarkable advances over the past two decades have enabled theoretical predictions of the band structure of semiconductors using fully ab initio methods,[5–8] but these approaches have not yet achieved a level of accuracy suitable for direct comparison with experiment. Furthermore, ab initio calculations 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 imposes additional limitations on the predictive power of ab initio calculations. By contrast, the goal of our work is transferability, simplicity, and accuracy, which can be obtained using analytical expressions with parameters fit to experimental data.

Starting from Kane's 8 ! 8 k · p model,[1,2,9] we will show that only a small number of parameters, especially the bandgap and one momentum matrix element,[10] 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, such as SiGeSn, GaAsSb, 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.[11] Our work is similar in some aspects, but we avoid the introduction of triple-index generalized Fermi–Dirac integrals.[12–14] Instead, we use Fermi–Dirac integrals, $F_j(\eta)$, that can be evaluated in MATLAB[15] using polylogarithm functions.[16] We discuss the validity of our approximations, present graphical representations of our results, and include detailed derivations as the supplementary material. 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. $E_n(\vec{k})$ is the energy of a band as a function of wave vector. 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). $\varepsilon_n(\vec{k})$ is the corresponding 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,[2,17] therefore, introduced a modified energy parameter,

$$E_n^u(\vec{k}) = E_n^u(\vec{k}) - \frac{\hbar^2 k^2}{2m_0}. \tag{2}$$

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 more recent years. The prime 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

The $\vec{k}\cdot\vec{p}$ electronic band structure method[18] takes advantage of the Bloch wave function $u_{n\vec{k}}(\vec{r}) = u_{n\vec{k}}(\vec{r})\exp(i\vec{k}\cdot\vec{r})$, 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 $Hu_{n0} = E_{n0}^u u_{n0}$ is known at the Γ-point for $\vec{k} = 0$ with wave functions $u_{n0}$ and eigenvalues $E_{n0}^u$, for example, from experimental measurements of the band energies. 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 be obtained by solving the eigenvalue problem,[1]

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

Here, $\langle n0|\vec{p}|i0\rangle$ 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, review articles,[9] and in the supplementary material. For small wave vectors $\vec{k}$, the solutions of (3) can be obtained using perturbation theory.

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 $|X + iY\rangle$, $|Z\rangle$ for the VB. Without loss of generality, we may assume that $\vec{k}$ points along the z-direction. The only nonvanishing 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. Including spin degeneracy, this yields an 8 × 8 matrix, with two identical 4 × 4 on-diagonal block matrices,[2,9]

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

and vanishing off-diagonal blocks. $\Delta_0$ 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.[10] More accurate $\vec{k}\cdot\vec{p}$-models 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).[19,20]

The matrix (4) has one obvious eigenvalue $E^u = 0$. 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,[17,21] 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.[22]

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

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

which can be solved analytically as described in the supplementary material 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 $\epsilon$.

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,[2,9]

$$\frac{1}{m^*_{lh}} = \frac{2E_P}{3E^u_0}, \quad (6)$$

$$\frac{1}{m^*_{so}} = \frac{E_P}{3(E^u_0 + \Delta_0)}, \quad (7)$$

$$\frac{1}{m^*_e} = 1 + \frac{E_P}{3}\left[\frac{2}{E^u_0} + \frac{1}{E^u_0 + \Delta_0}\right], \quad (8)$$

as shown in the supplementary material. Due to the large nonparabolicity of the bands, these effective mass values are only valid for very small values of k, as shown by the dotted lines in Fig. 1.

To obtain a simple analytical solution of Eq. (5), we use the large SO splitting approximation $E^u_0 \ll \Delta_0$. The characteristic equation (5) then becomes quadratic and offers solutions for the light-hole and electron bands,[2,9]

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

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 bandgap $E^u_0$, 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 equation (5) quite well, as the small difference between the solid and dashed lines in Fig. 1.

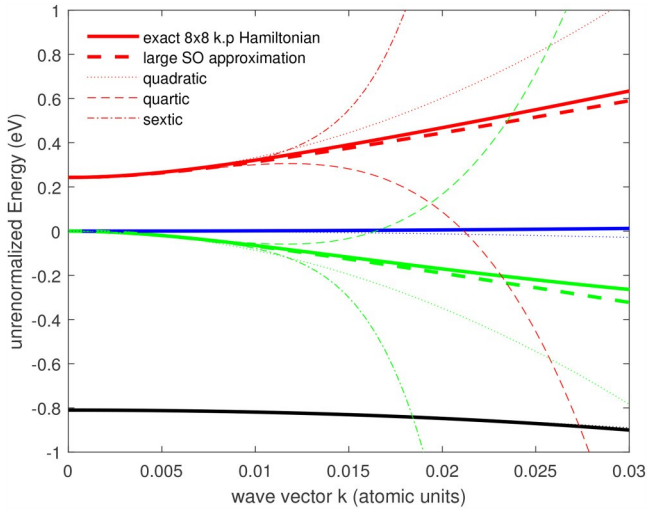


FIG. 1. Band structure for InSb at 0 K. Thick lines show the heavy hole (blue), light hole (green), split-off hole (black), and electron bands (red) from the cubic characteristic equation (5) (solid) and from the large SO approximation (9) (dashed) as a function of the wave vector k in atomic units (inverse 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 $k^6$ (dotted-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 the supplementary material for a similar graph showing the energies as a function of the square of the wave vector.

### 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[23,24]

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

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

for the CB and light-hole bands, respectively. We remind the reader that we introduced the excess energy $\epsilon$ in Sec. II A.

The nonparabolicity coefficients $\alpha_n$ and $\beta_n$ are defined by[23,24]

$$\frac{\hbar^2k^2}{2m_0m_n} = \epsilon^u_n\left(1 + \alpha_n\epsilon^u_n + \beta_n\epsilon^{u2}_n\right). \quad (15)$$

They can be obtained[23] by expanding Eqs. (13) and (14) into a power series of $\epsilon_n$,

$$\frac{\hbar^2k^2}{2m_0m_e\epsilon^u_e} = 1 + \frac{m^{*2}_e\epsilon^u_e}{4\mu^2_{lh}E^u_0}, \frac{m^{*4}_e\epsilon^{u2}_e}{4\mu^3_{lh}E^{u2}_0} \quad \text{and} \quad (16)$$

$$\frac{\hbar^2k^2}{2m_0m_{lh}\epsilon^u_{lh}} = 1 + \frac{m^{*2}_{lh}\epsilon^u_{lh}}{4\mu^2_{lh}E^u_0} + \frac{m^{*4}_{lh}\epsilon^{u2}_{lh}}{4\mu^3_{lh}E^{u2}_0}. \quad (17)$$

The nonparabolicity parameters in the large SO splitting approximation from an 8 ! 8 k · p-model are, therefore,[23–25]

$$\alpha_e = \frac{m_e^{*2}}{4\mu_{lh}^2 E_0^u} = \frac{\left(1 - m_e^{*}\right)^2}{E_0^u}, \tag{18}$$

$$\beta_e = -\frac{m_e^{*4}}{4\mu_{lh}^3 E_0^{u2}} = \frac{-\left(2m_e^{*}\right)\left(1 - m_e^{*}\right)^3}{E_0^{u2}}, \tag{19}$$

$$\alpha_{lh} = \frac{m_{lh}^{*2}}{4\mu_{lh}^2 E_0^u} = \frac{\left(1 + m_{lh}^{*}\right)^2}{E_0^u}, \tag{20}$$

$$\beta_{lh} = \frac{m_{lh}^{*4}}{4\mu_{lh}^3 E_0^{u2}} = \frac{2m_{lh}^{*}\left(1 + m_{lh}^{*}\right)^3}{E_0^{u2}}. \tag{21}$$

Figure 2 compares the "exact" solution of the 8 ! 8 k · p-Hamiltonian in the large SO splitting approximation given by Eqs.(13) and (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 ($\beta_h$=0) is about the same as the error caused by the large SO approximation. Adding the next term ($\beta_h$ = 0) makes the solution nearly indistinguishable from the exact large SO approximation. For this review, we, therefore, set $\beta_h = 0$ for applications of our nonparabolic band structure model.

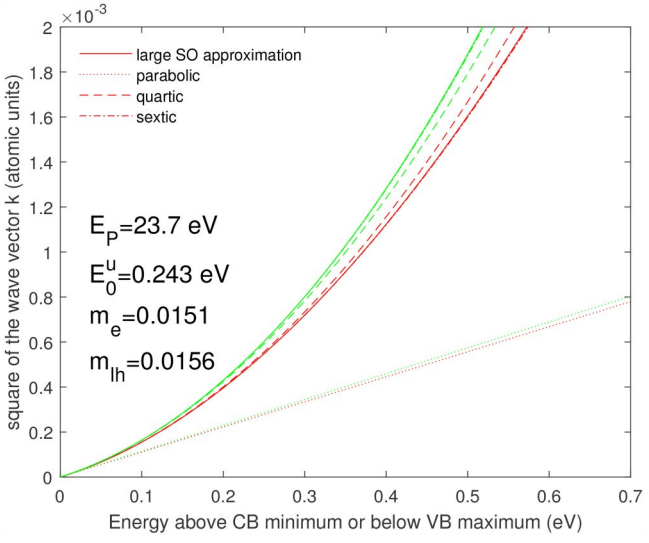


FIG. 2. The square of the wave vector (in atomic units) vs energy from the extremum for electrons (red) and light holes (green) in InSb at 0 K, calculated using Eqs. (13) and (14) within the large SO splitting approximation (solid), using parameters from Sec. II C. The dotted lines show the parabolic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

## D. Temperature dependence of the effective masses

In Secs. II A–II C, we did not address the temperature dependence of the band parameters that appear in the theoretical expressions. Since k · p-theory allows very accurate parameterizations of the band structure, incorporating temperature effects correctly is crucial for predicting the thermal properties as well as the results from optical measurements. To calculate the temperature dependence of the effective masses shown in Eqs. (10)–(12), we ignore the small variation of the momentum matrix element $E_P$ due to thermal expansion[26] and only consider the temperature dependence of the bandgap $E_0$. Furthermore, we do not consider the renormalization of the bandgap $E_0$ due to many-body effects. While these many-body effects play a role at higher temperatures and the corresponding high intrinsic carrier concentrations,[27,28] they are beyond the scope of the current paper and will be discussed elsewhere.

The principal cause of the temperature dependence of the electronic band structure is the electron-phonon interaction. Phonons are typically calculated within the "quasiharmonic" approximation. This consists of expanding the interatomic potential to quadratic order around an equilibrium position that depends on the temperature due to the thermal expansion of the lattice.[29] The origin of thermal expansion is the presence of nonzero anharmonic terms (higher than quadratic) in the expansion of the interatomic potential. Within the same scheme, the electronic band structure is calculated for a static lattice of atoms at the temperature-dependent equilibrium positions. Accordingly, all band parameters, including the direct bandgap, acquire a temperature dependence since the electronic eigenvalues are affected by the thermal variations of the lattice constant. We call this bandgap the "unrenormalized" bandgap $E_0^u(T)$.

However, this does not account for the full temperature dependence of the experimental bandgap $E_0^{exp}(T)$, i.e., the energy separation between the bottom of the CB and the top of the VB. (This is also called the "thermal gap," because it enters the calculation of the carrier concentration using Fermi–Dirac statistics.) The onset of the optical absorption, also known as the optical activation energy, may be higher than the experimental band gap due the Burstein–Moss shift.[30,31]) The dynamic deformations induced by the quasiharmonic lattice vibrations also affect the electronic band structure and renormalize the bandgap. Phenomenologically, the combined contributions from the electron-phonon interactions can be written for the particular case of $E_0$ as[32,33]

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

The first term describes the thermal expansion energy shift mentioned in the previous paragraph. The second term is the Debye–Waller contribution that 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. It is usually negative.[33] The third term is the

self-energy contribution that arises from the first-order electron-phonon Hamiltonian taken to second order in perturbation theory (emission or absorption of a phonon by an electron followed by reabsorption or re-emission, respectively) and is often positive.[33] The theory of this electron-phonon renormalization of the band gap has been described by Cardona and Gopalan.[32] An application of this theory to the direct bandgap of InSb was given in Ref. 33.

From the above description, it is apparent that a rigorous incorporation of electron-phonon effects into a $\vec{k}\cdot\vec{p}$ calculation requires first an evaluation of the unrenormalized band structure using temperature-dependent parameters, such as $E_0^u(T)$, which account for thermal expansion effects, followed by a calculation of the Debye–Waller and self-energy terms described by the last two terms in Eq. (22). Neither step is straightforward.

The thermal expansion term can be written[34] as

$$\left(\frac{\partial E_0}{\partial T}\right)_{TE} = \left(\frac{\partial E_0}{\partial T}\right)_p + (3\alpha B)\left(\frac{\partial E_0}{\partial p}\right)_T. \quad (23)$$

Here, $\alpha(T)$ is the temperature-dependent thermal expansion coefficient[35–39] and B the bulk modulus. Neglecting the temperature dependence of B = 46 GPa (taken from Ref. 40) and approximating the pressure derivative in Eq. (23) as the pressure derivative of the experimental bandgap at constant temperature (taken as 0.155 eV/GPa from Ref. 41), we obtain

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

The thermal expansion coefficient of zinc blende semiconductors is approximately given by[42]

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

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[15]), A is an adjustable parameter, and $\Theta_D$ the Debye temperature. The Debye temperature for InSb is about 168 K for InSb,[36] but we treat it as an adjustable parameter to fit the thermal expansion coefficients.[35] With parameters $A = 17.5 \times 10^{-6}\ \mathrm{K}^{-1}$ and $\Theta_D = 450$ K, satisfactory agreement with the experimental data can be achieved above 100 K; see Fig. 3. An ab initio calculation of the thermal expansion coefficient of InSb was performed by Miranda *et al.*[39]

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,[42]

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

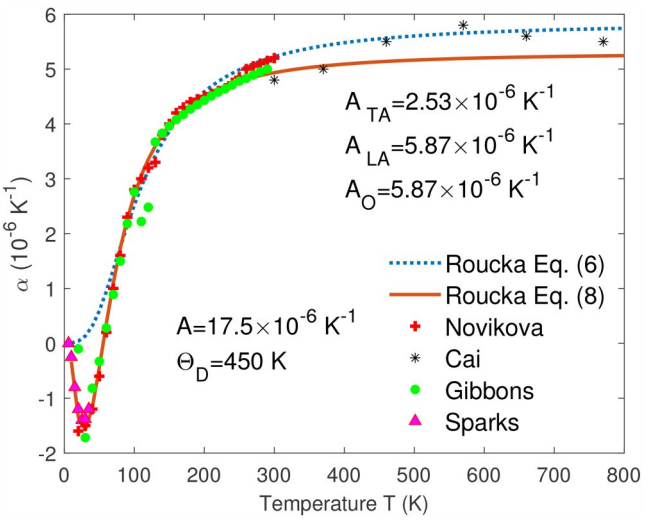


FIG. 3. Linear thermal expansion coefficient α vs temperature taken from the literature (Refs. 35–38) (symbols) along with a fit to the data using Eq. (25) (dotted) and Eq. (27) (solid).

The vibrational properties of InSb lead to phonon parameters $\Theta_{TA} = 53.7$ K, $\Theta_{LA} = 248$ K, and $\Theta_O = 244$ K, calculated as described in Ref. 42. The amplitudes are obtained as fit parameters: $A_{TA} = 2.53 \times 10^{-6}\ \mathrm{K}^{-1}$ and $A_{LA} = A_O = 5.87 \times 10^{-6}\ \mathrm{K}^{-1}$. The corresponding Grüneisen parameters[42] are $\gamma_{TA} = -0.86$ and $\gamma_{LA} = \gamma_O = 1.33$. Both expressions (25) and (27) lead to nearly the same thermal expansion shift of the bandgap (28) shown in Fig. 4; i.e., the negative thermal expansion coefficient at low temperatures is not a large contribution.

Unfortunately, the parameter $E_0(0\ \mathrm{K})$ in Eq. (24) is not accessible experimentally. It has been customary in the literature to use instead $E_0^{exp}(0\ \mathrm{K})$, but this is conceptually incorrect because, due to zero-point motion, the Debye–Waller and self-energy contributions in Eq. (22) do not vanish at zero temperature.

On the other hand, the calculation of the Debye–Waller and self-energy effects at many points in the Brillouin zone is computationally extremely costly and rarely performed. A possible solution would be using $E_0^{exp}(T)$ in $\vec{k}\cdot\vec{p}$-theory, but this assumes that the Debye–Waller and self-energy corrections for points k = 0 are fully determined by the corrections at k = 0 in the manner described by $\vec{k}\cdot\vec{p}$-theory, which is not justified.[43]

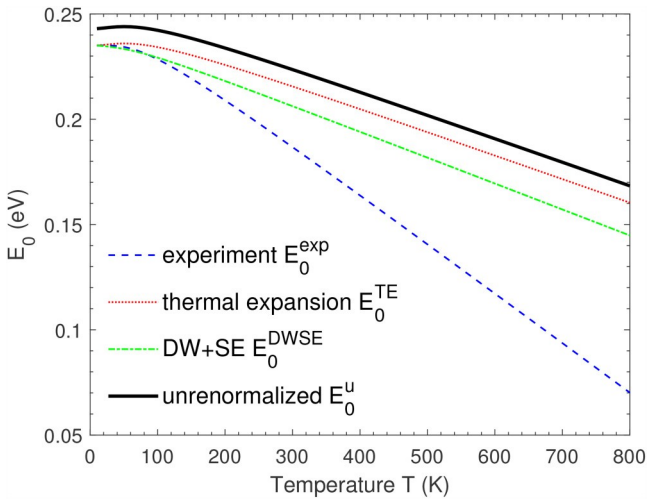


FIG. 4. Direct gap $E_0$ of InSb vs temperature. Dashed: experimental gap $E_0^{exp}$ from Eq. (28); dotted: thermal expansion contribution; dotted-dashed: Debye–Waller and self-energy contributions from Eq. (29); solid: unrenormalized bandgap from Eq. (24) for calculation of effective masses.

In the following, we describe an approximate way to obtain $E_0^u(0\ K)$ for application to the $k \cdot p$-theory. Combining this $E_0^u(0\ K)$ and the experimental effective masses, we obtain new values of the parameter $E_P$ (Sec. II F) that we subsequently (Sec. II) use to calculate important thermal properties, such as intrinsic carrier concentrations. The procedure involves the approximation that the Debye–Waller and self-energy corrections do not affect the effective mass. We find good agreement with experiment.

To find $E_0^u(0\ K)$, we proceed as follows: The experimental bandgap was determined to be[44]

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

with parameters $E_B^{exp} = 261$ meV (unrenormalized bandgap), $a_B^{exp} = 26$ meV (electron-phonon coupling strength), and $\Omega = 18.9$ meV (energy of the coupling phonon), correcting an error in Ref. 44). This result (28) overestimates the electron-phonon parameters, because it includes the redshift due to thermal expansion as well as due to the Debye–Waller (DW) and self-energy (SE) terms. To calculate the combined DW+SE shift

$$E_0^{DWSE}(T) = E_0^{exp}(0\ K) - 3B\left[\frac{\partial E_0^{exp}}{\partial p}\right]_{T=0} \alpha(\theta)\, d\theta, \quad (29)$$

we subtract the thermal expansion shift from the experimental bandgap and fit the difference (29) with a Bose–Einstein expression as given in Eq. (28). This results in parameters $E_B^{DWSE} = 243$ meV, $a_B^{DWSE} = 7.3$ meV, and $\Omega^{DWSE} = 10.9$ meV. By definition, the unrenormalized bandgap $E_0^u(0\ K)$ is equal to $E_B^{DWSE} = 243$ meV.

We are now able to calculate $E_0^u(T)$ with Eq. (24), which determines the $k \cdot p$-band structure.

The experimental bandgap $E_0^{exp}(T)$ and the contributions due to thermal expansion $E_0^{TE}(T)$ and DW+SE renormalization $E_0^{DWSE}(T)$ are shown in Fig. 4. The latter two are similar in magnitude over the complete range. About half of the redshift of the direct bandgap 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 $E_0^{TE}(T)$ but is shifted upward by 8 meV due to the renormalization of the low-temperature bandgap by zero-point phonon vibrations. (The electron-phonon shift obtained with the rigid pseudoion method was larger.[33]) At the lowest temperatures, a small (about 4 meV) increase of the band gap with increasing temperature can be seen due to the negative thermal expansion coefficient at low temperatures.

We are now able to calculate the temperature dependence of the effective electron and light-hole masses using Eqs. (10) and (11). Our results obtained from the $k \cdot p$-theory are shown in Fig. 5. At low temperatures, a small increase of the effective masses is seen due to the negative thermal expansion coefficient, followed by a decrease at higher temperatures. Experimentally, the temperature dependence of the effective electron mass of InSb from 40 to 260 K was found using magnetophonon magnetoresistance measurements by Stradling and Wood.[45] They found a 9% decrease of the bare electron mass from 40 to 260 K. This compares favorably with our calculated reduction by 6.6% over the same temperature range. The discrepancy could arise from experimental errors, the complicated theory for magnetoresistance, and the use of the large spin–orbit approximation in our calculations.

To summarize this section, we repeat that the unrenormalized (also known as mass) bandgap $E_0^u(T)$ given by Eq. (24) as shown in Fig. 4 must be used to calculate the effective masses in

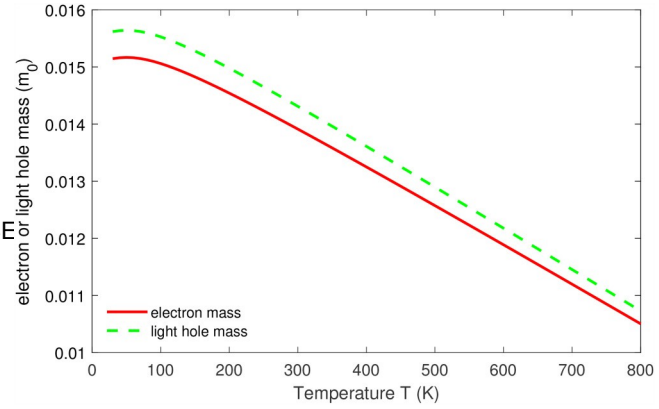


FIG. 5. Effective masses of the electron (solid) and light-hole (dashed) bands of InSb as a function of temperature, calculated taking into account only the thermal expansion contribution to the bandgap shift, the Debye–Waller and self-energy terms.

k·p-theory[11,45,46] using Eqs. (10)–(12), not the experimental bandgap (which is also known as the thermal bandgap).

### E. Weak polaron effects

In a weakly ionic crystal such as InSb, an electron or hole polarizes the ions and causes a small change of their equilibrium positions. As the carrier moves through the crystal, it must drag this ionic displacement with it. The carrier together with this strain field is called a polaron, which is described in many textbooks.[47,48] The mass of this polaron is measured in transport experiments, such as cyclotron resonance or magnetotransport. This polaron mass is larger than the "bare" electron or hole mass calculated with band structure methods, such as k·p-theory.

The ratio of the effective polaron mass $m_{pol}$ to the effective bare mass $m_{bare}$ is given by[47–52]

$$\frac{m_{pol}}{m_{bare}} \approx 1 + \frac{1}{6}\alpha_F, \tag{30}$$

where

$$\alpha_F = \frac{1}{2}\frac{e^2}{\hbar}\sqrt{\frac{2m_{bare}m_0}{E_{LO}}}\left(\frac{1}{\epsilon_1}-\frac{1}{\epsilon_s}\right) \tag{31}$$

is the Fröhlich coupling constant, $e$ the electronic charge, $E_{LO}$ the energy of the longitudinal optical (LO) phonon, and $\epsilon_1$ and $\epsilon_s$ the high-frequency and static dielectric constant, respectively. Using the Lyddane–Sachs–Teller relation for the longitudinal and transverse optical (TO) phonons,[53] we can write the term in parentheses as

$$\frac{1}{\epsilon_1}-\frac{1}{\epsilon_s} = \frac{1}{\epsilon_s}\left(\frac{E_{LO}^2}{E_{TO}^2}-1\right), \tag{32}$$

which is easier to evaluate using infrared spectroscopy.

For InSb, the Fröhlich coupling constant is only $\alpha_F \approx 0.022$, and the polaron correction is less than 1% smaller than the accuracy of the effective masses considered here. We are, therefore, justified to ignore polaron corrections for our purposes.

### F. Momentum matrix element and effective masses

It has been standard practice[11,46] to calculate k·p-band structures from the temperature-dependent bandgap $E_0^{TE}(T)$, shown by the dotted line in Fig. 4, that includes thermal expansion, but neglects the renormalization due to Debye–Waller and self-energy corrections at elevated temperatures. In other words, one first calculates the unrenormalized band structure in k·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 motion is included in $E_0^{TE}(T)$, while renormalization due to thermal excitation of phonons is not. We prefer an approach where the Debye–Waller and self-energy corrections are treated as a whole, given by the square brackets in Eq. (28). Therefore, the energy $E_0^u(T)$ should enter the k·p-band structure, not $E_0^{TE}(T)$.

This standard practice based on the use of $E_0^{TE}(T)$ to calculate effective masses has resulted in a comprehensive body of work, especially the compilation of matrix element parameters by Lawaetz for many different semiconductors.[10] Lawaetz calculated k·p-parameters based on cyclotron measurements of the effective masses and experimental low-temperature bandgaps, which include the renormalization due to zero-point phonon motion. To exclude all Debye–Waller and self-energy corrections 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 bandgap $E_0^u = 0.243$ eV, we solve Eq. (6) to obtain

$$E_P = \frac{3}{2}E_0^u\left(1+\frac{1}{m_{lh}^*}\right) = 23.7 \text{ eV}, \tag{33}$$

somewhat larger than the usual value of 23.1 eV published by Lawaetz.[10] The corresponding effective electron mass (8) with $\Delta_0 = 0.81$ eV equals $m_e^* = 0.0136$, which is in excellent agreement with the experimental value. If we instead use the expressions (10) and (11) from the large SO splitting approximation, then the light-hole mass remains the same, but the effective electron mass increases to 0.0151, slightly larger than the experimental value of 0.014. 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,[53]

$$g_n(\epsilon_n) = \frac{1}{4\pi^3}\int d^3k\,\delta\left(E_{n\vec{k}}-\epsilon_n\right) = \frac{1}{\pi^2}\int_0^\infty k^2 dk\,\delta\left(E_{nk}-\epsilon_n\right). \tag{34}$$

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[24]

$$dk = \sqrt{\frac{m_0 m_n}{2\hbar^2\epsilon_n}}\frac{1+2\alpha_n\epsilon_n+3\beta_n\epsilon_n^2}{\sqrt{(1+\alpha_n\epsilon_n+\beta_n\epsilon_n^2)(1+2\alpha_n\epsilon_n+3\beta_n\epsilon_n^2)}}\,d\epsilon_n \tag{35}$$

and

$$k^2 dk = \frac{1}{2}\left(\frac{2m_0 m_n}{\hbar^2}\right)^{\frac{3}{2}}\sqrt{\epsilon_n(1+\alpha_n\epsilon_n+\beta_n\epsilon_n^2)(1+2\alpha_n\epsilon_n+3\beta_n\epsilon_n^2)}\,d\epsilon_n. \tag{36}$$

The density of states is, therefore,[11,53]

$$g_n(\epsilon_n) = \frac{1}{2\pi^2}\left(\frac{2m_0 m_n^*}{\hbar^2}\right)^{\frac{3}{2}}\sqrt{\epsilon_n(1+\alpha_n\epsilon_n+\beta_n\epsilon_n^2)(1+2\alpha_n\epsilon_n+3\beta_n\epsilon_n^2)}: \tag{37}$$

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

$$\left(1 + 2\alpha_n\epsilon_n + 3\beta_n\epsilon_n^2\right)\sqrt{1 + \alpha_n\epsilon_n + \beta_n\epsilon_n^2} \approx 1 + \frac{5}{2}\alpha_n\epsilon_n, \quad (38)$$

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

$$m^*_{n,\mathrm{DOS}}(\epsilon_n) = m_n^*\left(1 + \alpha_n\epsilon_n + \beta_n\epsilon_n^2\right)^{1/3}\left(1 + 2\alpha_n\epsilon_n + 3\beta_n\epsilon_n^2\right)^{2/3}. \quad (39)$$

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

$$m^*_{e,\mathrm{DOS}}(\epsilon) \approx m_e^*\left(1 + \frac{5}{3}\alpha_e\epsilon_e\right). \quad (40)$$

In other words, when the excess energy is equal to the bandgap $E_0$ (i.e., $\alpha_e\epsilon_e = 1$), 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 the $\epsilon_e$ term (shown by the dotted line). We, therefore, have confidence that the expansion (15) converges well in the approximation for large SO splittings. Since $\alpha_e\epsilon_e$ is not exactly small, one might wonder to what extent the linear expansion (40) is accurate. As shown in Fig. 6, the linearization of Eq. (40) 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 "8-band model" or with the large SO approximation. The linearization (40) of the density of states is necessary to evaluate the chemical potential of a degenerate electron gas using Fermi–Dirac integrals.

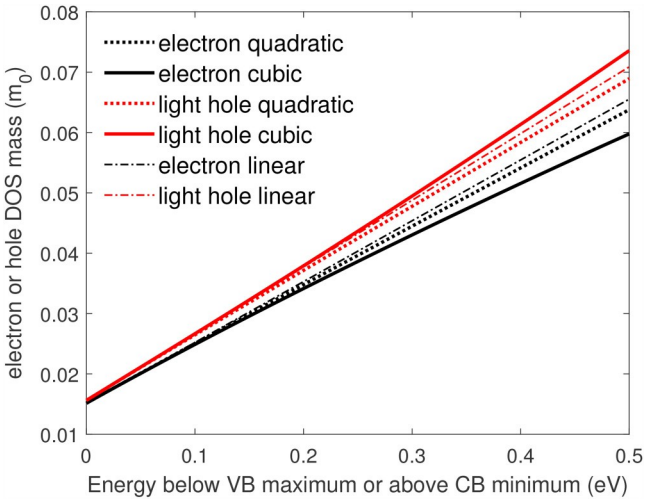


FIG. 6. Effective density-of-states electron mass $m^*_{e,\mathrm{DOS}}$ (black) and light-hole mass $m^*_{lh,\mathrm{DOS}}$ (red) of InSb at 0 K as a function of excess energy above the conduction band minimum or below the valence band maximum, calculated from Eq. (39) (solid). The dotted lines show the results with $\beta_n = 0$. The dashed-dotted lines show the linear expansion (40).

## B. Chemical potential and intrinsic carrier concentration vs temperature

We apply the density of states (34) for nonparabolic bands to calculate the chemical potential $\mu$ and the intrinsic carrier concentration $n$ for InSb as a function of temperature $T$.

The electron density $n_\Gamma$ in the $\Gamma$-valley of the CB at temperature $T$ is[53,54]

$$n_\Gamma(T) = \int_0^\infty d\epsilon\, g_e(\epsilon) f\left(E_0^{\exp}(T) + \epsilon\right), \quad \text{where} \quad (41)$$

$$f(E) = \left[\exp\left(\frac{E-\mu}{k_BT}\right) + 1\right]^{-1} \quad (42)$$

is the Fermi–Dirac distribution function with the chemical potential $\mu$ and the Boltzmann constant $k_B$. Note that we use the experimental (or "thermal") bandgap $E_0^{\exp}(T)$ in Eq. (41), not the "mass" bandgap $E_0(T)$ introduced in Sec. II D.

By setting $\beta_n = 0$ and keeping only terms linear in $\alpha_n$, we find that the density-of-states enhancement factor (38) is approximately $1 + \frac{5}{2}\alpha_e\epsilon_e$. With the substitutions $y = \epsilon/k_BT$ and $x = (\mu - E_0^{\exp})/k_BT$, the electron density can be written using Fermi–Dirac integrals as[55,56]

$$n_\Gamma(T) = N_e(T)\left[F_{\frac{1}{2}}\left(\frac{\mu - E_0^{\exp}}{k_BT}\right) + \frac{15}{4}\alpha_e k_BT F_{\frac{3}{2}}\left(\frac{\mu - E_0^{\exp}}{k_BT}\right)\right], \quad (43)$$

with the prefactor[53,54]

$$N_n(T) = \frac{1}{4}\left(\frac{2m_0 m_n^* k_BT}{\pi\hbar^2}\right)^{3/2}. \quad (44)$$

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[56,57]

$$p_{lh}(T) = N_{lh}(T)\left[F_{\frac{1}{2}}\left(-\frac{\mu}{k_BT}\right) + \frac{15}{4}\alpha_{lh}k_BT F_{\frac{3}{2}}\left(-\frac{\mu}{k_BT}\right)\right]. \quad (45)$$

For the heavy hole band we do not consider the nonparabolicity and set $\alpha_{hh} = 0$. We fix the heavy hole mass at $m_{hh} = 0.43$, independent of temperature.[24] This mass is determined by the separation $E_0'$ between the p-bonding VB and the p-antibonding CB at the $\Gamma$-point, which has a weak relative temperature dependence.[58] This

results in[56]

$$p_{hh}(T) = N_{hh}(T)F_{\frac{1}{2}}\left(-\frac{\mu}{k_B T}\right), \quad (46)$$

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,[56]

$$p_{so}(T) = N_{so}(T)F_{\frac{1}{2}}\left(-\frac{\Delta_0 - \mu}{k_B T}\right), \quad (47)$$

$$n_L(T) = 4N_L(T)F_{\frac{1}{2}}\left(\frac{\mu - E_L^{exp}}{k_B T}\right), \quad (48)$$

$$n_X(T) = 3N_X(T)F_{\frac{1}{2}}\left(\frac{\mu - E_X^{exp}}{k_B T}\right), \quad (49)$$

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.[1,10] The dominant contribution to $m_{so}^*$ comes from the spin–orbit splitting $\Delta_0$, and therefore, the smaller gap $E_0$ 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 $E_0^{exp}$. There are four L-valleys and three X-valleys in zinc blende semiconductors. (Diamond-type semiconductors have six $\Delta$-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,[46] as far as we know. Percentages of the electron and hole populations in the various bands are shown in the supplementary material.

## IV. RESULTS AND DISCUSSION

We can find the chemical potential of an intrinsic semiconductor from the charge neutrality condition,[59]

$$n_\Gamma(T) + n_L(T) + n_X(T) = p_{lh}(T) + p_{hh}(T) + p_{so}(T), \quad (50)$$

at a given temperature T, for example, using polylogarithm functions[16,44] in MATLAB.[15] 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" bandgap $E_0^{exp} = 0.187$ eV and the unrenormalized mass bandgap $E_0^u = 0.221$ eV. The room-temperature effective masses are $m_e^* = 0.0138$ and $m_{lh}^* = 0.0142$, calculated using Eqs. (10) and (11) in the large SO splitting approximation. For the holes, the light-hole density is only a very small contribution, because the heavy

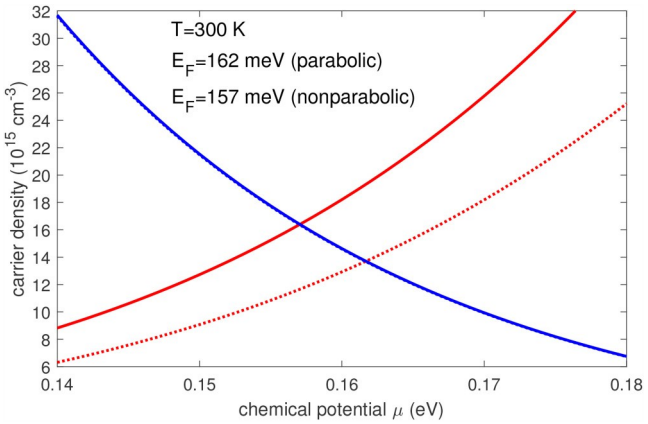


FIG. 7. Electron (red) and hole density (blue) of InSb as a function of chemical potential at 300 K in the parabolic approximation (dotted) and with the lowest nonparabolic corrections (solid). The thermal and mass bandgaps were taken from Fig. 4, and the electron and light-hole masses were calculated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

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. (39), and therefore, the electron density is larger than in the parabolic case, because the prefactor Eq. (44) is proportional to $m_e^{*1.5}$.

For a given temperature T, we plot n and p as a function of $\mu$. The intrinsic chemical potential is found at the location where the two lines cross,[60] thus satisfying the charge neutrality condition (50). In the parabolic case, the electron and hole densities vs 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 $E_0^{exp} = 0.187$ eV as mentioned earlier. The intrinsic carrier concentration of InSb at 300 K is $13.6 \times 10^{15}$ cm$^{-3}$ for parabolic bands and $16.4 \times 10^{15}$ 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[11]). At low temperatures, the chemical potential is approximately equal to half the band gap,[53] and therefore, the argument of the Fermi integral is very small. For this case, we can apply the nondegenerate limit,[61–63]

$$F_{\frac{1}{2}}(\eta) \approx \exp(\eta) \quad \text{for} \quad \eta \ll -1, \quad (51)$$

essentially using classical Maxwell–Boltzmann statistics to describe the electron and hole populations. This approximation leads to the well-known expressions,[53,54]

$$\mu \approx \frac{E_0}{2} + \frac{3}{4}k_B T \ln\left(\frac{m_{hh}^*}{m_e^*}\right) \quad \text{and} \quad (52)$$

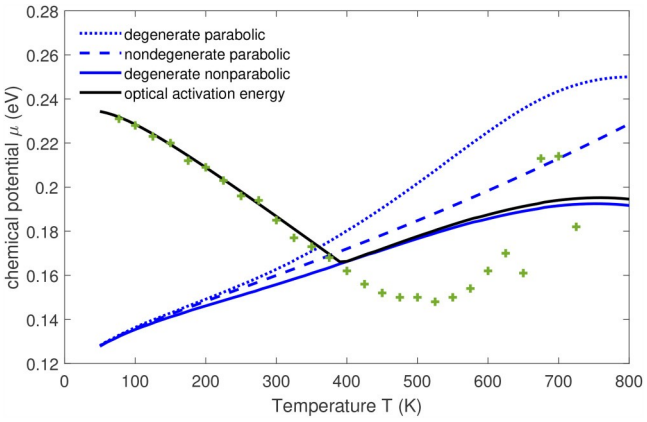


FIG. 8. Chemical potential vs temperature for parabolic bands in the nondegenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest nonparabolic correction in the large SO approximation. The full temperature dependence of the direct gap according to Eq. (28) was included in the Fermi–Dirac integral, effective masses were calculated taking into account only the thermal expansion contribution to the bandgap given by Eq. (24), not the Debye–Waller and self-energy corrections. The optical activation energy from (53) and the experimental direct bandgap (Ref. 44) from a fit to the temperature-dependent ellipsometrically determined dielectric function with a Johs–Herzinger parametric oscillator (symbols) are also shown.

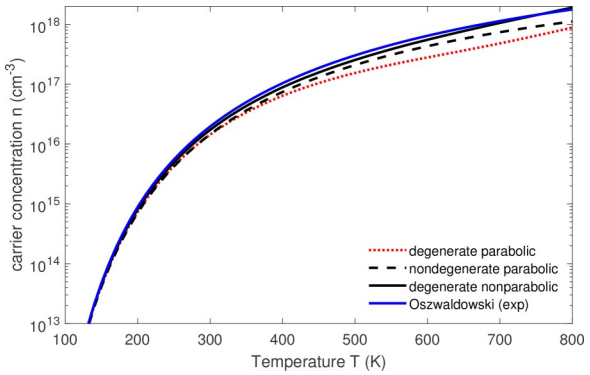


FIG. 9. Intrinsic carrier concentration vs temperature for parabolic bands in the nondegenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest parabolic correction in the large SO approximation. The temperature dependence of the direct gap according 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 bandgap from (24). The blue line shows a fit to carrier concentrations determined from Hall measurements by Oszwaldowski and Zimpel (Ref. 46).

$$n = 2\left(\frac{m_0 k_B T}{2\pi\hbar^2}\right)^{3/2}(m_e^* m_{hh}^*)^{3/4}\exp\left(-\frac{E_0}{2k_B T}\right). \quad (53)$$

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. (52). 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.[16] 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 nondegenerate (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.[60] If nonparabolic bands are considered, then the effective electron mass becomes larger, which increases the carrier concentration according to Eq. (53). Oszwaldowski and Zimpel[46] obtained the temperature dependence of the intrinsic carrier concentration of InSb from 200 to 800 K with Hall measurements. Assuming a Hall scattering factor of unity, they found an intrinsic carrier concentration near $1.8 \times 10^{18}\,\mathrm{cm}^{-3}$ at 800 K. They fitted their results with the expression[46]

$$n = 2.9 \times 10^{11}(2400 \cdot T)^{0.75}\left(1 + 2.7 \times 10^{-4} \cdot T \cdot T^{1.5}\right) \times \exp\left(-\frac{0.129 \cdot 1.5 \times 10^{-4} \cdot T}{k_B T}\right), \quad (54)$$

where n is in units of $\mathrm{cm}^{-3}$, 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 $1.9 \times 10^{18}\,\mathrm{cm}^{-3}$ 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. (53), 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.[46]

It is a common practice in the interpretation of low-field transport measurements of semiconductors to assume that the Hall scattering factor

$$r_H = \frac{\hbar\langle\tau^2\rangle}{\hbar\langle\tau\rangle^2} \quad (55)$$

is unity, where $\tau$ is the scattering time and $\hbar\langle\tau\rangle$ indicates the energy weighted average within the carrier population of the band. This assumption is not always true, however, and deviations from unity by up to 10%–100% are common.[64–67] The value of the Hall scattering factor varies based on the particular type of scattering process and its dependence on the energy of the carriers. A complicated dependence of $r_H$ on temperature and doping concentration is often found.

We also compare the results for the chemical potential in Fig. 8 with optical measurements of the band gap.[44] The optical activation energy (i.e., the bandgap observed in optical absorption or an ellipsometry experiment) is increased through the Burstein–Moss shift and given by[68]

$$E_A = \max\left(E_0, E_0 + \left(1 + \frac{m_e^*}{m_{hh}^*}\right)(\mu - E_0)\right). \quad (56)$$

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 bandgap. 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.[44] shown by symbols at higher temperatures.

## V. SUMMARY

We have shown how a simple $8 \times 8$ $k \cdot p$ 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 bandgap $E_0$ and the spin–orbit splitting $\Delta_0$ 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[46] if the unrenormalized bandgap $E_0^u$ 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 Debye–Waller and self-energy electron-phonon interactions. The replacement of the experimental bandgap $E_0$ by the unrenormalized gap $E_0^u$ requires a small adjustment of the momentum matrix element.

## SUPPLEMENTARY MATERIAL

The following content is provided as supplementary material: (1) a detailed derivation of the $k \cdot p$ band structure of InSb and related materials (GaAs, α-Sn) based on Kane's $8 \times 8$ Hamiltonian[2] from degenerate perturbation theory with analytical and graphical results in various approximations; (2) a survey of different treatments for the nonparabolicity of the electron and light-hole bands and the resulting electron and hole density of states; (3) general expressions for the chemical potential and intrinsic carrier concentration of semiconductors with nonparabolic bands and their evaluation based on several scenarios for the temperature dependence of the effective mass; (4) a discussion of the linear thermal expansion coefficient of InSb based on the model of Roucka et al.;[42] and (5) the occupation of the various electron and hole bands of InSb at elevated temperatures.

## 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 No. FA9550-20-1-0135 and FA9550-24-1-0061. This material is based upon work supported by the National Science Foundation under Award No. DMR-2235447. S.Z. and C.A.A. are grateful to Arnold Kiefer and Bruce Claflin for their hospitality at the AFRL Sensors Directorate in Dayton, OH, where part of this work was performed. This material is based on research sponsored by the Air Force Research Laboratory (AFRL) under Agreement No. FA9453-23-2-0001. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory (AFRL) or the U.S. Government.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

Stefan Zollner: Conceptualization (lead); Formal analysis (lead); Funding acquisition (lead); Investigation (lead); Methodology (lead); Project administration (lead); Software (lead); Supervision (lead); Validation (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (lead). Carlos A. Armenta: Formal analysis (equal); Methodology (equal); Software (equal); Visualization (equal). Sonam Yadav: Formal analysis (equal); Investigation (equal); Software (equal); Visualization (equal). José Menéndez: Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Methodology (equal); Writing – review & editing (equal).

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES

[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, Oxford, 2010).
[5] M. L. Cohen and J. R. Chelikowsky, Electronic Structure and Optical Properties of Semiconductors (Springer, Berlin, 2012).
[6] J. R. Chelikowsky and M. L. Cohen, J. Appl. Phys. 117, 112812 (2015).
[7] F. Bechstedt, F. Fuchs, and G. Kresse, Phys. Status Solidi B 246, 1877 (2009).
[8] M. Rohlfing and S. G. Louie, Phys. Rev. B 62, 4927 (2000).

[9]E. O. Kane, in Semiconductors and Semimetals, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1966), Vol. 1, p. 75.
[10]P. Lawaetz, Phys. Rev. B 4, 3460 (1971).
[11]R. A. Masut, Eur. J. Phys. 43, 015501 (2022).
[12]W. Zawadzki and J. Kolodziejczak, Phys. Status Solidi 6, 409 (1964).
[13]W. Zawadzki, R. Kowalczyk, and J. Kolodziejczak, Phys. Status Solidi 10, 513 (1965).
[14]W. Zawadzki, Adv. Phys. 23, 435 (1974).
[15]See http://www.mathworks.com.
[16]M. D. Ulrich, W. F. Seng, and P. A. Barnes, J. Comput. Electron. 1, 431 (2002).
[17]E. O. Kane, J. Phys. Chem. Solids 1, 82 (1956).
[18]W. Shockley, Phys. Rev. 78, 173 (1950).
[19]M. Cardona and F. H. Pollak, Phys. Rev. 142, 530 (1966).
[20]D. Rideau, M. Feraille, L. Ciampolini, M. Minondo, C. Tavernier, H. Jaouen, and A. Ghetti, Phys. Rev. B 74, 195208 (2006).
[21]G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. 98, 368 (1955).
[22]R. W. Cunningham and J. B. Gruber, J. Appl. Phys. 41, 1804 (1970).
[23]E. M. Conwell and M. O. Vassell, Phys. Rev. 166, 797 (1968). Note that this reference addresses the small SO limit, $\Delta E_0$.
[24]F. J. Bartoli, J. R. Meyer, C. A. Hoffmann, and R. E. Allen, Phys. Rev. B 27, 2248 (1983).
[25]C. Jacoboni and L. Reggiani, Rev. Mod. Phys. 55, 645 (1983).
[26]C. Emminger, N. S. Samarasingha, M. Rivero Arias, F. Abadizaman, J. Menendez, and S. Zollner, J. Appl. Phys. 131, 165701 (2022).
[27]P. Vashishta and R. K. Kalia, Phys. Rev. B 25, 6492 (1982).
[28]R. Zimmermann, Phys. Status Solidi B 146, 371 (1988).
[29]M. S. Haque, Phys. Rev. B 12, 1501 (1975).
[30]E. Burstein, Phys. Rev. 93, 632 (1954).
[31]T. S. Moss, Proc. Phys. Soc. B 67, 775 (1954).
[32]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.
[33]S. Zollner, S. Gopalan, and M. Cardona, Solid State Commun. 77, 485 (1991).
[34]S. Zollner, M. Garriga, J. Humlíček, S. Gopalan, and M. Cardona, Phys. Rev. B 43, 4349 (1991).
[35]S. I. Novikova, in Semiconductors and Semimetals, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1966), Vol. 2, p. 33.
[36]X. Cai and J. Wei, J. Appl. Phys. 114, 083507 (2013).
[37]D. F. Gibbons, Phys. Rev. 112, 136 (1958).
[38]P. W. Sparks and C. A. Swenson, Phys. Rev. 163, 779 (1967).
[39]A. L. Miranda, B. Xu, O. Hellman, A. D. Romero, and M. J. Verstraete, Semicond. Sci. Technol. 29, 124002 (2014).
[40]P. E. Van Camp, V. E. Van Doren, and J. T. Devreese, Phys. Rev. B 41, 1598 (1990).
[41]I. M. Booth, M. H. Hawton, and W. J. Keeler, Phys. Rev. B 25, 7713 (1982).
[42]R. Roucka, Y.-Y. Fang, J. Kouvetakis, A. V. G. Chizmeshya, and J. Menéndez, Phys. Rev. B 81, 245214 (2010).
[43]S. Gopalan, P. Lautenschlager, and M. Cardona, Phys. Rev. B 35, 5577 (1987).
[44]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).
[45]R. A. Stradling and P. A. Wood, J. Phys. C 3, L94 (1970).
[46]M. Oszwaldowski and M. Zimpel, J. Phys. Chem. Solids 49, 1179 (1988).
[47]C. Kittel, Introduction to Solid State Physics (Wiley, New York, 2004).
[48]M. Grundmann, The Physics of Semiconductors (Springer International Publishing, Cham, Switzerland, 2021).
[49]D. Emin, Polarons (Cambridge University, Cambridge, 2013).
[50]J. Appel, Solid State Phys. 21, 193 (1968).
[51]R. P. Feynman, Phys. Rev. 97, 660 (1955).
[52]N. Bouarissa, S. A. Siddiqui, and M. Boucenna, Cryst. Res. Technol. 52, 1700018 (2017).
[53]N. W. Ashcroft and N. D Mermin, Solid State Physics (Saunders, Philadelphia, PA, 1976).
[54]S. M. Sze, Physics of Semiconductor Devices (Wiley, New York, 1981).
[55]H. Ehrenreich, Phys. Rev. 120, 1951 (1960).
[56]J. Menéndez, C. D. Poweleit, and S. E. Tilton, Phys. Rev. B 101, 195204 (2020).
[57]C. A. Hoffman, J. R. Meyer, R. J. Wagner, F. J. Bartoli, M. A. Engelhardt, and H. Höchst, Phys. Rev. B 40, 11693 (1989).
[58]S. Logothetidis, L. Viña, and M. Cardona, Phys. Rev. B 31, 947 (1985).
[59]J. S. Blakemore, Semiconductor Statistics (Pergamon, Oxford, 1962).
[60]W. Shockley, Electrons and Holes in Semiconductors (Van Nostrand, Princeton, NJ, 1950).
[61]J. McDougall and S. E. Clifton, Philos. Trans. R. Soc. London Ser. A 237, 67 (1938).
[62]J. S. Blakemore, Solid-State Electron. 25, 1067 (1982).
[63]R. Kim, X. Wang, and M. Lundstrom, Notes on Fermi-Dirac Integrals, 3rd ed. (2008); see https://nanohub.org/resources/5475.
[64]J. Menéndez, C. Xu, and J. Kouvetakis, Mater. Sci. Semicond. Process. 164, 107596 (2023).
[65]D. A. Anderson and N. Apsley, Semicond. Sci. Tech. 1, 187 (1987).
[66]N. Apsley, D. A. Anderson, and J. B. Morrison, Semicond. Sci. Tech. 2, 44 (1987).
[67]D. L. Rode, Semicond. Semimetals 10, 1 (1975).
[68]P. K. Chakraborty, G. C. Datta, and K. P. Ghatak, Phys. B 339, 198 (2003).