Atomistic k·p theory

Craig E. Pryor and M.-E. Pistol

Articles you may be interested in

Deriving k·p parameters from full-Brillouin-zone descriptions: A finite-element envelope function model for quantum-confined wurtzite nanostructures
J. Appl. Phys. **116**, 033709 (2014); 10.1063/1.4890585

Atomistic calculations of Ga(NAsP)/GaP(N) quantum wells on silicon substrate: Band structure and optical gain
Appl. Phys. Lett. **100**, 111901 (2012); 10.1063/1.3694028

Atomistic pseudopotential theory of optical properties of exciton complexes in InAs/InP quantum dots
Appl. Phys. Lett. **99**, 231106 (2011); 10.1063/1.3669507

Type II heterostructures formed by zinc-blende inclusions in InP and GaAs wurtzite nanowires
Appl. Phys. Lett. **97**, 041910 (2010); 10.1063/1.3473779

Band alignment between GaAs and partially ordered GaInP
Appl. Phys. Lett. **80**, 3111 (2002); 10.1063/1.1472478

Atomistic $k \cdot p$ theory

Craig E. Pryor[1,a)] and M.-E. Pistol[2,b)]

[1]*Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa 52242, USA*
[2]*NanoLund and Solid State Physics, Lund University, P.O. Box 118, 221 00 Lund, Sweden*

Pseudopotentials, tight-binding models, and $k \cdot p$ theory have stood for many years as the standard techniques for computing electronic states in crystalline solids. Here, we present the first new method in decades, which we call atomistic $k \cdot p$ theory. In its usual formulation, $k \cdot p$ theory has the advantage of depending on parameters that are directly related to experimentally measured quantities, however, it is insensitive to the locations of individual atoms. We construct an atomistic $k \cdot p$ theory by defining envelope functions on a grid matching the crystal lattice. The model parameters are matrix elements which are obtained from experimental results or *ab initio* wave functions in a simple way. This is in contrast to the other atomistic approaches in which parameters are fit to reproduce a desired dispersion and are not expressible in terms of fundamental quantities. This fitting is often very difficult. We illustrate our method by constructing a four-band atomistic model for a diamond/zincblende crystal and show that it is equivalent to the sp^3 tight-binding model. We can thus directly derive the parameters in the sp^3 tight-binding model from experimental data. We then take the atomistic limit of the widely used eight-band Kane model and compute the band structures for all III–V semiconductors not containing nitrogen or boron using parameters fit to experimental data. Our new approach extends $k \cdot p$ theory to problems in which atomistic precision is required, such as impurities, alloys, polytypes, and interfaces. It also provides a new approach to multiscale modeling by allowing continuum and atomistic $k \cdot p$ models to be combined in the same system. © *2015 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4936170]

I. INTRODUCTION

An electron in the periodic potential of a semiconductor may be described using pseudopotentials,[1,2] tight-binding models,[3–6] or $k \cdot p$ theory.[7] All three have been applied to semiconductor nanostructures[5,8–16] in which the translational symmetry is broken by heterostructures, an applied potential, or a finite size. Pseudopotentials and tight-binding models are inherently atomistic in that they allow, and even require, specification of the locations of atoms. In contrast, $k \cdot p$ theory provides a Hamiltonian for the coarse grained crystal using parameters that depend on the composition and structure of the material. Alloys are treated in the virtual crystal approximation in which the parameters specifying the band structure are empirically fit to the observed electronic properties of a material. While this precludes a description with atomic scale precision, typically one does not know the exact position of every atom in a system anyway.

The three methods involve tradeoffs in the approximations made and the physical phenomena that they describe most accurately. While $k \cdot p$ theory is a continuum model, the momentum matrix elements which parameterize it depend on the atomic scale structure of the electronic wave functions. This is advantageous in the computation of optical properties, since the dipole matrix elements depend on the momentum matrix elements of the Bloch functions which also determine the band structure. Tight-binding models use atomistic scale wave functions but involve a large number of parameters which are

determined using complicated fitting procedures[17] such as genetic algorithms.[5,18] Pseudopotentials also require a large number of form factors and must rely on complex fitting procedures, especially when strain is involved.[19] Pseudopotentials are atomistic, but smooth out the core wave function which results in smaller momentum matrix elements and thus smaller optical matrix elements. By including enough bands, any of the three methods can be made to be accurate throughout the Brillouin zone.[20–24] Here, we will focus on the dispersion around zone center.

$k \cdot p$ theory in the envelope approximation has been used successfully to describe electronic states in a wide variety of inhomogeneous semiconductor systems. The electronic wave function is taken to be a sum of Bloch functions, each multiplied by a slowly varying envelope function. The effective Hamiltonian for the envelopes consists of material-dependent coefficients multiplying derivatives acting on the envelopes. The electronic states are then determined by putting the envelopes on a computational grid and using finite difference approximations for derivatives, giving a model that is coarse-grained over a size comparable to the grid spacing.

An interesting question arises: can the computational grid be made small enough to make the model atomistic? Finite differences make the model superficially resemble a tight-binding model with hopping between grid sites, suggesting a connection between tight-binding and $k \cdot p$ models. In this paper, we will develop an atomistic $k \cdot p$ theory by constructing it on a grid in which the sites correspond to atomic positions in the crystal lattice. We will show that the

a)Electronic mail: craig-pryor@uiowa.edu
b)Electronic mail: mats-erik.pistol@ftf.lth.se

atomistic limit of a simple four-band $k \cdot p$ theory is equivalent to a tight-binding model, thus allowing the determination of $k \cdot p$ parameters and tight-binding parameters in terms of each other. This connection may be used to derive atomistic models that identically reproduce the long wavelength physics of $k \cdot p$ theory.

We will begin in Sec. II with a discussion of $k \cdot p$ theory in the envelope approximation (henceforth simply $k \cdot p$ theory) on a three-dimensional grid of points, and using finite differences. We generalize this widely used method to an arbitrary grid which need not be cartesian or regular. In Sec. III, we take the grid to be the crystal lattice itself and introduce difference operators on a diamond or zincblende grid. In Sec. IV, we discuss how matrix elements are computed on such an atomistic grid and examine the new non-zero momentum matrix elements which appear. In Sec. V, we take the atomistic limit of a simple four-band model without spin-orbit coupling and show that it is identical to a tight-binding model. In Sec. VI, we examine the non-Hermiticity that can arise (as in the four-band model) and in Sec. VII show how it may be resolved using the finite volume method. In Sec. VIII, we take the atomistic limit of the more realistic (and widely used) eight-band Kane model with spin-orbit coupling. In Sec. IX, we describe our fitting method and present our numerical fits of the atomistic parameters for the non-nitride III–V semiconductors. We conclude with a discussion of some of the unique features and merits of the atomistic limit.

II. ENVELOPE THEORY

We begin with the basics in order to establish notation. The Hamiltonian for a single electron in a semiconductor is

$$\hat{H} = \hat{H}_0 + \hat{H}_{so}, \tag{1a}$$

$$\hat{H}_0 = \frac{\hat{p}^2}{2m_0} + V_0(\mathbf{r}) + V_e(\mathbf{r}), \tag{1b}$$

$$\hat{H}_{so} = \frac{\hbar}{4m_0^2 c^2}(\sigma \times \nabla V) \cdot \hat{\mathbf{p}}, \tag{1c}$$

where $V_0(\mathbf{r})$ is the crystal potential, and $V_e(\mathbf{r})$ is a possible externally applied potential. In a translational invariant system, the single electron states may be found using multiband $k \cdot p$ theory by writing the wave function as

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sum_{n=1}^{N} a_{n\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r}) u_n(\mathbf{r}), \tag{2}$$

where $u_n(\mathbf{r})$ are the zone-center Bloch functions, and $a_{n\mathbf{k}}$ are numerical coefficients. The Hamiltonian is then an $N \times N$ numerical matrix

$$H_{mn} = -\frac{\hbar^2}{2m_0}k^2 \, \delta_{mn} - \frac{i\hbar}{m_0}\mathbf{k} \cdot \langle u_m | \hat{\mathbf{p}} | u_n \rangle$$
$$+ \langle u_m | \hat{H}_0 | u_n \rangle + \left\langle u_m \left| \frac{\hbar^2}{4m_0^2 c^2}(\sigma \times \nabla V) \right| u_n \right\rangle \cdot \mathbf{k}, \tag{3}$$

where N is the number of zone-center Bloch functions included in the basis. We will initially omit the spin-orbit

term and restore it later in Section VIII. Note that there is no explicit requirement of small $\mathbf{k}$ and Eq. (3) is exact except for errors introduced by truncating the basis. For any finite value of N, the accuracy decreases with increasing $\mathbf{k}$, since the solution will not be accurately expressed as a finite sum of zone center Bloch functions.

If translation symmetry is broken due to a heterojunction or an applied potential then the plane waves of Eq. (2) are replaced with envelope functions, and the wave function becomes

$$\psi(\mathbf{r}) = \sum_{n=1}^{N} f_n(\mathbf{r}) u_n(\mathbf{r}), \tag{4}$$

where the $u_n(\mathbf{r})$ are Bloch functions, but otherwise quite arbitrary. Substituting $\psi(\mathbf{r})$ into the Schrödinger equation, and multiplying both sides by $u_m^*(\mathbf{r})$, we obtain

$$\left[-\frac{\hbar^2}{2m_0}a_{mn}(\mathbf{r})\nabla^2 + \mathbf{P}_{mn}(\mathbf{r}) \cdot \nabla + v_{mn}(\mathbf{r}) \right] f_n(\mathbf{r})$$
$$= E a_{mn}(\mathbf{r}) f_n(\mathbf{r}), \tag{5a}$$

$$a_{mn}(\mathbf{r}) = u_m^*(\mathbf{r}) u_n(\mathbf{r}), \tag{5b}$$

$$\mathbf{P}_{mn}(\mathbf{r}) = \frac{\hbar}{im_0} u_m^*(\mathbf{r}) \hat{\mathbf{p}} u_n(\mathbf{r}), \tag{5c}$$

$$v_{mn}(\mathbf{r}) = u_m^*(\mathbf{r}) \left[\frac{1}{2m_0}\hat{\mathbf{p}}^2 + V(\mathbf{r}) \right] u_n(\mathbf{r}), \tag{5d}$$

where $\mathbf{P}_{mn}$ and v_{mn} are simply functions (i.e., the operators they contain act only on the Bloch functions contained within them). We may make one re-arrangement which will be useful when we consider Hermiticity

$$\left[-\frac{\hbar^2}{2m_0}\nabla a_{mn}(\mathbf{r})\nabla + \mathcal{P}_{mn}(\mathbf{r}) \cdot \nabla + v_{mn}(\mathbf{r}) \right] f_n(\mathbf{r})$$
$$= E a_{mn}(\mathbf{r}) f_n(\mathbf{r}), \tag{6a}$$

$$\mathcal{P}_{mn}(\mathbf{r}) = \frac{1}{2}\left(\mathbf{P}_{mn} - \mathbf{P}_{nm}^* \right). \tag{6b}$$

Since u_m may be chosen as real, $\mathbf{P}_{mn}$ is real, and therefore $\mathcal{P}_{mn} = -\mathcal{P}_{nm}$. The above equations give the effective Schrödinger equation for the envelope functions in which the Bloch functions appear as parameters. Note that both are exact even with an incomplete set of Bloch functions, since the envelopes are arbitrary. For example, the trivial case of a single Bloch function $u_1(\mathbf{r}) = 1$ simply gives back the original Schrödinger equation, in which case the solution would consist of an envelope with variation over atomic scales. If a sufficiently large Bloch basis is used, however, the wave function is well approximated with slowly varying envelopes. Eq. (5a) will be approximate if the f_n are constrained, as they are when defined on a grid that imposes a momentum cutoff.

To obtain numerical solutions for systems that are not amenable to analytic methods requires reduction to a discrete system, such as by using a finite basis set or functions

defined on a grid. For nanostructures with irregular geometries, putting the envelope functions on a grid and replacing derivatives with finite differences is especially convenient, since no assumptions about the geometric symmetry are required.[13,25,26] We denote the grid points with coordinates $\mathbf{R}$ and use $\mathbf{r}$ for continuum coordinates. The continuous space can be broken up into cells $\Omega_{\mathbf{R}}$, each centered on the grid site at $\mathbf{R}$, and the integral over all space can then be written as a sum of integrals over the cells

$$\int d^3r = \sum_{\mathbf{R}} \int_{\Omega_{\mathbf{R}}} d^3r. \tag{7}$$

If a sufficiently large number of Bloch functions are used the envelope functions will be slowly varying, and $f_n(\mathbf{r})$ and its derivatives will be approximately constant over each cell. This seemingly reasonable assumption can result in a non-Hermitian Hamiltonian, which will be resolved in Sec. VI. Integrating Eq. (5a) over $\Omega_{\mathbf{R}}$ and approximating derivatives of $f_n(\mathbf{r})$ by finite differences on the grid, we obtain

$$\left[-\frac{\hbar^2}{2m_0} \langle u_m|u_n\rangle_{\Omega_{\mathbf{R}}} \Delta^2 + \frac{\hbar}{im_0} \langle u_m|\hat{\mathbf{p}}|u_n\rangle_{\Omega_{\mathbf{R}}} \cdot \Delta \right.$$
$$\left. + \langle u_m|\hat{H}_0|u_n\rangle_{\Omega_{\mathbf{R}}} \right] f_{n\mathbf{R}} = E\langle u_m|u_n\rangle_{\Omega_{\mathbf{R}}} f_{n\mathbf{R}}, \tag{8}$$

where $f_{n\mathbf{R}}$ is the nth envelope function on the site at $\mathbf{R}$, and Δ is the finite difference approximation to the gradient, $\Delta f_{n\mathbf{R}} \approx \partial_x f_n(\mathbf{r})|_{\mathbf{r}=\mathbf{R}}$, which is a weighted sum of the values of f_n at $\mathbf{R}$ and nearby grid sites. We adopt an abbreviated notation for the projected matrix element of an operator $\hat{\mathcal{O}}$

$$\int_{\Omega_{\mathbf{R}}} u_m^*(\mathbf{r})\hat{\mathcal{O}}u_n(\mathbf{r})\, d^3r = \langle u_m|\hat{\mathcal{O}}|u_n\rangle_{\Omega_{\mathbf{R}}}. \tag{9}$$

If $\Omega_{\mathbf{R}}$ contains an integer number of crystal unit cells then $\langle u_m|u_n\rangle_{\Omega_{\mathbf{R}}} = \delta_{mn}$. The solution of Eq. (8) is obtained by computing the eigenvalues and eigenvectors of a large sparse matrix, for which there are efficient algorithms.[27,28]

   If the Bloch functions are the same throughout the structure, then the $\langle u_m|\hat{\mathbf{p}}|u_n\rangle_{\Omega_{\mathbf{R}}}$ are constants and Eq. (8) can be used directly to determine the electronic states. This will be the case if confinement is provided by an externally applied potential or for a nanocrystal in which the vacuum is modeled as a large potential barrier. In a heterostructure, however, the matrix elements appearing in Eq. (8) will vary spatially. In the atomistic limit, in which the $\mathbf{R}$s correspond to individual atoms, the matrix elements will vary spatially even in a bulk crystal if it contains different atoms. This will cause the Hamiltonian in Eq. (8) to be non-Hermitian, requiring a more careful treatment. We will return to this problem and its remedy in Section VI.

III. ATOMISTIC GRID

   The finite difference approximation consists of replacing a derivative at a grid site with a difference operator which acts by taking weighted sums of the values on nearby grid sites. For example, on a uniform cartesian grid, the derivative of a function f at a grid site located at $\mathbf{R}$ may be approximated using the symmetric difference

$$\Delta_x f(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{R}} = \frac{f(\mathbf{R}+\epsilon\hat{\mathbf{x}}) - f(\mathbf{R}-\epsilon\hat{\mathbf{x}})}{2\epsilon} = \partial_x f(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{R}} + \mathcal{O}(\epsilon), \tag{10}$$

where ϵ is the grid spacing, and $\hat{\mathbf{x}}$ is a unit vector. This gives the lowest order approximation to ∂_x, and more accurate results may be obtained by including more sites in the sum.[29] The accuracy of the finite difference approximation improves as the grid spacing decreases, making it tempting to shrink the grid spacing as much as possible. The existence of a physical crystal lattice suggests using the crystal lattice itself as the computational grid. The values of the envelope functions will then be defined on the atoms themselves, yielding an atomistic theory. We will develop this model for the diamond/zincblende lattice because of its importance in semiconductor physics, but the approach can be applied to any crystal structure.

   On a rectangular grid, the low order finite difference approximations may be written down intuitively, but on a non-rectilinear grid one needs a more systematic approach. The general method for constructing a difference operator is to write down a Taylor series expansion of a function at a point $\mathbf{R}$, express the function values on sites in terms of that expansion, and solve for the linear combination of the function values on the site and its neighbors that gives the desired derivative to lowest order.[29–32] This method is used to obtain high-order difference approximations with smaller errors, and it may be used with non-rectilinear grids or even irregular grids.

   Because the zincblende/diamond lattice has a basis with two atoms per unit cell, the difference operators will be different on the inequivalent sites. We denote the atom at $(0, 0, 0)$ as type 1 (assumed to be the anion in zincblende) and the atom at $\frac{a_{latt}}{4}(1, 1, 1)$ as type 2 (cation in zincblende), where a_{latt} is the lattice constant. The nearest neighbors of the type 1 atoms are at displacements

$$\mathbf{d}_1 = \frac{a_{latt}}{4}(1, 1, 1), \tag{11a}$$

$$\mathbf{d}_2 = \frac{a_{latt}}{4}(-1, -1, 1), \tag{11b}$$

$$\mathbf{d}_3 = \frac{a_{latt}}{4}(-1, 1, -1), \tag{11c}$$

$$\mathbf{d}_4 = \frac{a_{latt}}{4}(1, -1, -1), \tag{11d}$$

and the nearest neighbors of the type 2 atoms are at displacements $-\mathbf{d}_n$. On the type 1 sites, the nearest neighbor differences are given by

$$\partial_x f(\mathbf{r})|_{\mathbf{r}=\mathbf{R}} = \Delta_{1x} f(\mathbf{R}) + \mathcal{O}(a_{latt})$$
$$= [f(\mathbf{R}+\mathbf{d}_1) - f(\mathbf{R}+\mathbf{d}_2) - f(\mathbf{R}+\mathbf{d}_3)$$
$$+ f(\mathbf{R}+\mathbf{d}_4)]/a_{latt} + \mathcal{O}(a_{latt}), \tag{12a}$$

$$\partial_y f(\mathbf{r})|_{\mathbf{r}=\mathbf{R}} = \Delta_{1y}f(\mathbf{R}) + \mathcal{O}(a_{latt})$$
$$= [f(\mathbf{R}+\mathbf{d}_1) - f(\mathbf{R}+\mathbf{d}_2) + f(\mathbf{R}+\mathbf{d}_3)$$
$$- f(\mathbf{R}+\mathbf{d}_4)]/a_{latt} + \mathcal{O}(a_{latt}), \quad (12b)$$

$$\partial_z f(\mathbf{r})|_{\mathbf{r}=\mathbf{R}} = \Delta_{1z}f(\mathbf{R}) + \mathcal{O}(a_{latt})$$
$$= [f(\mathbf{R}+\mathbf{d}_1) + f(\mathbf{R}+\mathbf{d}_2) - f(\mathbf{R}+\mathbf{d}_3)$$
$$- f(\mathbf{R}+\mathbf{d}_4)]/a_{latt} + \mathcal{O}(a_{latt}), \quad (12c)$$

$$\nabla^2 f(\mathbf{r})|_{\mathbf{r}=\mathbf{R}} = \Delta_1^2 f(\mathbf{R}) + \mathcal{O}(a_{latt})$$
$$= \frac{8}{a_{latt}^2}[f(\mathbf{R}+\mathbf{d}_1) + f(\mathbf{R}+\mathbf{d}_2) + f(\mathbf{R}+\mathbf{d}_3)$$
$$+ f(\mathbf{R}+\mathbf{d}_4) - 4f(\mathbf{R})] + \mathcal{O}(a_{latt}), \quad (12d)$$

and on the type 2 sites the difference operators are

$$\Delta_{2x}f(\mathbf{R}) = [-f(\mathbf{R}-\mathbf{d}_1) + f(\mathbf{R}-\mathbf{d}_2) + f(\mathbf{R}-\mathbf{d}_3)$$
$$- f(\mathbf{R}-\mathbf{d}_4)]/a_{latt}, \quad (13a)$$

$$\Delta_{2y}f(\mathbf{R}) = [-f(\mathbf{R}-\mathbf{d}_1) + f(\mathbf{R}-\mathbf{d}_2) - f(\mathbf{R}-\mathbf{d}_3)$$
$$+ f(\mathbf{R}-\mathbf{d}_4)]/a_{latt}, \quad (13b)$$

$$\Delta_{2z}f(\mathbf{R}) = [-f(\mathbf{R}-\mathbf{d}_1) - f(\mathbf{R}-\mathbf{d}_2) + f(\mathbf{R}-\mathbf{d}_3)$$
$$+ f(\mathbf{R}-\mathbf{d}_4)]/a_{latt}, \quad (13c)$$

$$\Delta_2^2 f(\mathbf{R}) = \frac{8}{a_{latt}^2}[f(\mathbf{R}-\mathbf{d}_1) + f(\mathbf{R}-\mathbf{d}_2) + f(\mathbf{R}-\mathbf{d}_3)$$
$$+ f(\mathbf{R}-\mathbf{d}_4) - 4f(\mathbf{R})]. \quad (13d)$$

Using the site at $\mathbf{R}$ and its four nearest neighbors, only the four derivatives ∇ and ∇^2 can be constructed. Substituting the above difference operators in Eq. (8) gives the Hamiltonian for the envelopes, parameterized by the matrix elements in Eq. (8) which would be empirically fit to measurements on bulk materials.

Perturbative $k \cdot p$ models require approximations for second derivatives as well. Since no derivative approximation beyond ∇ and ∇^2 can be constructed with four nearest neighbors, second nearest neighbors must be used. The second nearest neighbor differences are the same for type 1 and 2 sites

$$\partial_x^2 f(\mathbf{R}) = \frac{1}{a_{latt}^2}[4f(\mathbf{R}) - f(\mathbf{R}+\mathbf{d}_{-1,-1,0}) - f(\mathbf{R}+\mathbf{d}_{1,-1,0})$$
$$- f(\mathbf{R}+\mathbf{d}_{-1,1,0}) - f(\mathbf{R}+\mathbf{d}_{1,1,0})$$
$$- f(\mathbf{R}+\mathbf{d}_{-1,0,-1}) - f(\mathbf{R}+\mathbf{d}_{1,0,-1})$$
$$- f(\mathbf{R}+\mathbf{d}_{-1,0,1}) - f(\mathbf{R}+\mathbf{d}_{1,0,1})$$
$$+ f(\mathbf{R}+\mathbf{d}_{0,-1,-1}) + f(\mathbf{R}+\mathbf{d}_{0,1,-1})$$
$$+ f(\mathbf{R}+\mathbf{d}_{0,-1,1}) + f(\mathbf{R}+\mathbf{d}_{0,1,1})] + \mathcal{O}(a_{latt}^2), \quad (14a)$$

$$\partial_x \partial_y f(\mathbf{x})|_{\mathbf{x}=\mathbf{R}} = \frac{1}{a_{latt}^2}[f(\mathbf{R}+\mathbf{d}_{1,-1,0}) + f(\mathbf{R}+\mathbf{d}_{-1,1,0})$$
$$- f(\mathbf{R}+\mathbf{d}_{1,1,0}) - f(\mathbf{R}+\mathbf{d}_{-1,-1,0})] + \mathcal{O}(a_{latt}^2), \quad (14b)$$

where $\mathbf{d}_{i,j,k} = \frac{a_{latt}}{2}[i\hat{\mathbf{x}} + j\hat{\mathbf{y}} + k\hat{\mathbf{z}}]$ is the displacement to the second nearest neighbor from the central site at $\mathbf{R}$.

Difference formulas for ∂_y^2, ∂_z^2, $\partial_x\partial_x$, and $\partial_y\partial_z$ are obtained by cyclic permutation of x, y, z. Note that while the approximation to ∇^2 involves only nearest neighbors, mixed derivatives and ∂_x^2, ∂_y^2, and ∂_z^2 require next nearest neighbors. This does not present any fundamental or technical problems and is consistent with the source of these terms, second order perturbation theory. Since the zeroth-order Hamiltonian contains nearest neighbor couplings from ∇^2, the second order Hamiltonian contains next nearest neighbor couplings.

IV. MATRIX ELEMENTS

The atomistic matrix elements are somewhat different from those usually appearing in $k \cdot p$ theory due to the projection of the Bloch states to atomistic cells. The most obvious choice for $\Omega_{\mathbf{R}}$ would be the Wigner Seitz cell around each $\mathbf{R}$, as shown in Fig. 1 for the diamond/zincblende lattice. Many of the same selection rules from $k \cdot p$ theory apply since they depend on T_d symmetry, which the atomic cells possess. The most notable difference is the existence of matrix elements that are zero in the continuum model but are non-zero in the atomistic limit with cancellations between the atomistic cells. Over the primitive unit cell, $\Omega_1 + \Omega_2$, the Bloch functions satisfy

$$\langle u_m | u_n \rangle_{\Omega_1+\Omega_2} = \langle u_m | u_n \rangle_{\Omega_1} + \langle u_m | u_n \rangle_{\Omega_2} = \delta_{mn}, \quad (15)$$

but the atomistic matrix elements are not necessarily proportional to δ_{mn}, since for $m \neq n$ there could be two nonzero terms that cancel. If no two bands transform as the same representation of T_d, then $\langle u_m | u_n \rangle_{\Omega_{1,2}} \propto \delta_{mn}$. For example, in a model with an $SXYZ$ basis, $\langle X | S \rangle_{\Omega_{1,2}} = 0$ because under a rotation by π about the y-axis the crystal is invariant, but
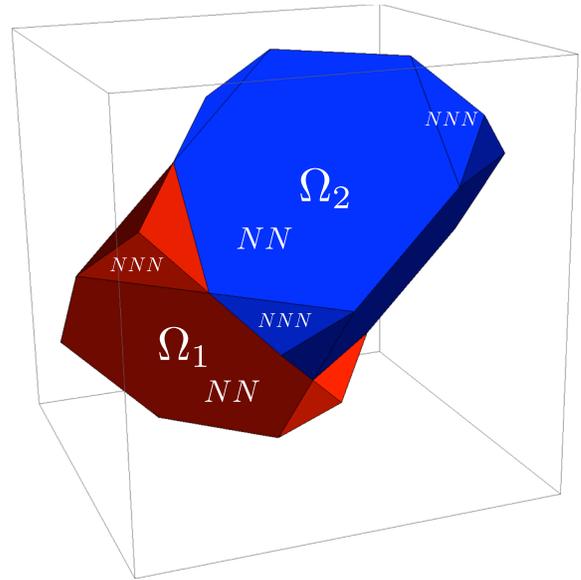
FIG. 1. Wigner Seitz cells in a zincblende crystal. Ω_1 is the cell around the type-I atom at $(0, 0, 0)$, and Ω_2 is the cell around the type-II atom at $(1/4, 1/4, 1/4)$. The hexagonal faces are the planes separating nearest neighbors, and the small triangular faces are the planes separating second nearest neighbors, which are of the same type. If only the nearest neighbor planes were included, the cells would be tetrahedra. Taking next nearest neighbors into consideration truncates the corners of the tetrahedra, replacing them with (shorter) triangular pyramids.

$\langle X|S\rangle_{\Omega_{1,2}}$ will change sign. In this paper, we will consider models in which all the states transform differently, and in particular, models derived from an $SXYZ$ basis. In these cases the right hand side of Eq. (8) is a diagonal matrix, which we choose to be a multiple of the unit matrix, and there is no need to solve a generalized eigenvalue problem. It is only necessary to multiply the left hand side by the inverse of this matrix.

The use of atomistic cells modifies the momentum matrix elements as well. For states projected to a volume Ω, the momentum matrix element is given by

$$\int_{\Omega} u_m^*(\mathbf{r}) \frac{\hbar}{i} \nabla u_n(\mathbf{r}) \, d^3r = \left(\int_{\Omega} u_n^*(\mathbf{r}) \frac{\hbar}{i} \nabla u_m(\mathbf{r}) \, d^3r \right)^* + \int_{\Omega} \frac{\hbar}{i} \nabla \left[u_m^*(\mathbf{r}) u_n(\mathbf{r}) \right] d^3r, \quad (16)$$

where the second integral on the right hand side will vanish due to periodicity if Ω contains an integer number of unit cells. If Ω contains some fraction of a unit cell, then $u_m^*(\mathbf{r}) u_n(\mathbf{r})$ is not periodic over Ω and there is no reason for the second integral on the right hand side to vanish. We write the projected matrix element in the more compact form

$$\langle u_m | \hat{\mathbf{p}} | u_n \rangle_{\Omega_i} = \langle u_n | \hat{\mathbf{p}} | u_m \rangle_{\Omega_i} + \Pi_{\Omega_i nm}, \quad (17)$$

where $\Pi_{\Omega_i nm}$ is given by the second integral on the right hand side of Eq. (16) and is nonzero only if Ω does not contain an integer number of unit cells. If we sum over two subcells Ω_1 and Ω_2 to make a whole unit cell, the correction must vanish, and therefore $\Pi_{\Omega_1 nm} = -\Pi_{\Omega_2 nm}$.

The atomistic momentum matrix elements between the conduction and valence bands obey the same selection rules as in continuum $k \cdot p$ theory, since they rely only on T_d symmetry, however because the projected matrix elements on the two atoms can be different, we have

$$iP_0 = \frac{\hbar}{m_0} \langle S | \hat{p}_x | X \rangle_{\Omega_1 + \Omega_2} = \frac{\hbar}{m_0} \langle S | \hat{p}_x | X \rangle_{\Omega_1} + \frac{\hbar}{m_0} \langle S | \hat{p}_x | X \rangle_{\Omega_2} = iP_{a_1} + iP_{a_2}, \quad (18)$$

where the subscript a denotes that the matrix elements are projected to single atoms. Throughout this paper, we will use an a subscript to distinguish atomistic parameters from those of the continuum $k \cdot p$ theory. For the diamond crystal $P_{a_1} = P_{a_2}$ due to inversion symmetry, while for zincblende $P_{a_1} \neq P_{a_2}$.

In $k \cdot p$ models with more than one p-like band, such as the 16 and 14-band models,[33,34] there are also matrix elements of the form $iQ = \frac{\hbar}{m_0} \langle X_v | \hat{p}_y | Z_c \rangle$ where the v and c subscripts indicate the valence and conduction bands. Matrix elements with this general form but within the same band, such as $\langle X_v | \hat{p}_y | Z_v \rangle$, are zero by a simple symmetry argument which depends on the periodicity of the unit cell. Due to invariance under a 180° rotation about the y-axis, $\langle X_v | \hat{p}_y | Z_v \rangle = \langle Z_v | \hat{p}_y | X_v \rangle$. For matrix elements over a whole unit cell, we also have $\langle X_v | \hat{p}_y | Z_v \rangle = \langle Z_v | \hat{p}_y | X_v \rangle^*$, and therefore $\langle X_v | \hat{p}_y | Z_v \rangle = \langle X_v | \hat{p}_y | Z_v \rangle^*$. But since the states $|X\rangle$, $|Y\rangle$, $|Z\rangle$ can all be taken as real, $\langle X_v | \hat{p}_y | Z_v \rangle$ must be pure imaginary and therefore the matrix element is zero. In the atomistic

case, we see from Eq. (17) that $\langle X_v | \hat{p}_y | Z_v \rangle_{\Omega_i} \neq \langle X_v | \hat{p}_y | Z_v \rangle_{\Omega_i}^*$ and the above argument breaks down. While the projected matrix element can be nonzero, the matrix element over the whole cell is zero and therefore

$$0 = \frac{\hbar}{m_0} \langle X_v | \hat{p}_y | Z_v \rangle_{\Omega_1} + \frac{\hbar}{m_0} \langle X_v | \hat{p}_y | Z_v \rangle_{\Omega_2} = iQ_a - iQ_a. \quad (19)$$

We see that the atomistic $k \cdot p$ model has an additional parameter not present in the continuum $k \cdot p$ model from which it is derived. In an inversion non-symmetric crystal, the atomistic limit will also double the number of Hamiltonian matrix elements since there will be different matrix elements on each atom.

Summarizing, in the atomistic model, we have

$$iP_{a_i} = \frac{\hbar}{m_0} \langle S | \hat{p}_x | X \rangle_{\Omega_i} = \frac{\hbar}{m_0} \langle S | \hat{p}_y | Y \rangle_{\Omega_i} = \frac{\hbar}{m_0} \langle S | \hat{p}_z | Z \rangle_{\Omega_i}, \quad (20a)$$

$$iP'_{a_i} = -\frac{\hbar}{m_0} \langle X | \hat{p}_x | S \rangle_{\Omega_i} = -\frac{\hbar}{m_0} \langle Y | \hat{p}_y | S \rangle_{\Omega_i} = -\frac{\hbar}{m_0} \langle Z | \hat{p}_z | S \rangle_{\Omega_i}, \quad (20b)$$

$$iP_{a_1} + iP_{a_2} = -iP'_{a_1} - iP'_{a_2} = iP_0, \quad (20c)$$

where P_0 is the usual continuum $k \cdot p$ parameter. In addition, there are new intraband matrix elements

$$iQ_{a_i} = \frac{\hbar}{m_0} \langle X | \hat{p}_y | Z \rangle_{\Omega_i} = \frac{\hbar}{m_0} \langle Z | \hat{p}_x | Y \rangle_{\Omega_i} = \frac{\hbar}{m_0} \langle Y | \hat{p}_z | X \rangle_{\Omega_i}$$
$$= \frac{\hbar}{m_0} \langle Z | \hat{p}_y | X \rangle_{\Omega_i} = \frac{\hbar}{m_0} \langle X | \hat{p}_z | Y \rangle_{\Omega_i} = \frac{\hbar}{m_0} \langle Y | \hat{p}_x | Z \rangle_{\Omega_i}, \quad (21)$$

which satisfy $iQ_{a_1} = -iQ_{a_2}$.

As will be seen in Sec. V, in an inversion non-symmetric crystal, the combination of finite differences and the fact that $P_i \neq P'_i$ results in a non-Hermitian Hamiltonian. One solution is to simply use an inversion symmetric basis, which is reasonable since $k \cdot p$ theory is often formulated in the symmetric approximation. As will be shown in Sec. VIII, inversion symmetry may still be broken by the sub-unit cell structure of the envelope function. Alternatively, we may change the volumes of Ω_1 and Ω_2 by using generalized Voronoi cells.[35,36] By adjusting Ω_1 and Ω_2, we can make $P_i = P'_i$ while maintaining inversion non-symmetry. A generalized Voronoi cell may be constructed by rescaling the distances that would be used to determine the Wigner Seitz cell. Consider a site at $\mathbf{R}$, with nearest neighbors at $\mathbf{R}_{NN,i}$, and next nearest neighbors at $\mathbf{R}_{NNN,j}$. The generalized Voronoi cell around $\mathbf{R}$ is the set of points $\mathbf{p}$ satisfying the two conditions

$$a|\mathbf{p} - \mathbf{R}| < |\mathbf{p} - \mathbf{R}_{NN,i}|, \quad (22a)$$

$$|\mathbf{p} - \mathbf{R}| < |\mathbf{p} - \mathbf{R}_{NNN,j}|, \quad (22b)$$

where a is a scaling factor that determines the relative sizes of a cell. Note that the second condition does not have a

scaling factor because the next nearest neighbors are of the same type. Using $a \neq 1$ on the type-1 sites, and $a \to 1/a$ on the type-2 sites, we change the relative volumes of Ω_1 and Ω_2 while maintaining $\Omega_1 + \Omega_2$ as a unit cell. Shrinking Ω_1 will decrease both P_{a_1} and P'_{a_1} while increasing P_{a_2} and P'_{a_2}, and therefore we may adjust the scaling factor to make $P_{a_1} = P'_{a_2}$ and $P'_{a_1} = P_{a_2}$. This method will prove useful for restoring Hermiticity in the inversion non-symmetric case in Sec. V.

V. FOUR-BAND MODEL

To demonstrate the basic structure of the atomistic limit, we first consider the simple four-band $k \cdot p$ model without spin orbit coupling, using zone-center Bloch states $|X\rangle$, $|Y\rangle$, $|Z\rangle$ for the valence band and $|S\rangle$ for the conduction band. This model and the tight-binding model with which it will be compared are too simple to be used for realistic calculations, but they demonstrate the basic structure of the atomistic limit. Using plane waves for the envelopes, the $k \cdot p$ Hamiltonian is

$$
H_0 = \begin{pmatrix}
E_c + \mathcal{E}_c(k) & iP_0 k_x & iP_0 k_y & iP_0 k_z \\
-iP_0 k_x & E_v + \mathcal{E}_v(k) & 0 & 0 \\
-iP_0 k_y & 0 & E_v + \mathcal{E}_v(k) & 0 \\
-iP_0 k_z & 0 & 0 & E_v + \mathcal{E}_v(k)
\end{pmatrix},
$$

$$\tag{23a}$$

with column labels $S \quad X \quad Y \quad Z$.

$$
iP_0 = \frac{\hbar}{m_0} \langle S|\hat{p}_x|X\rangle, \tag{23b}
$$

$$
\mathcal{E}_c(k) = \left(\frac{1}{2} + F_c\right)\frac{\hbar^2}{m_0}k^2, \tag{23c}
$$

$$
\mathcal{E}_v(k) = \left(\frac{1}{2} + F_v\right)\frac{\hbar^2}{m_0}k^2, \tag{23d}
$$

where F_c and F_v are remote band contributions to the conduction and valence band, respectively, and are included to make the $k \cdot p$ model agree with a tight-binding model. In the atomistic limit, the Hamiltonian also includes terms involving the momentum matrix element of Eq. (21)

$$
H_Q = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & iQ_a k_z & iQ_a k_y \\
0 & iQ_a k_z & 0 & iQ_a k_x \\
0 & iQ_a k_y & iQ_a k_x & 0
\end{pmatrix},
$$

$$\tag{24a}$$

with column labels $S \quad X \quad Y \quad Z$.

$$
iQ_a = \frac{\hbar}{m_0}\langle X|\hat{p}_y|Z\rangle_{\Omega_1} = -\frac{\hbar}{m_0}\langle X|\hat{p}_y|Z\rangle_{\Omega_2}. \tag{24b}
$$

Eq. (24a) is only for an anion site, and on the cation site we will have $-H_Q$. In a plane wave basis, it is convenient to define the approximate wave vectors obtained from the finite difference operators acting on plane waves

$$
\mathcal{K}_{1n} = -ie^{-i\mathbf{k}\cdot\mathbf{r}}\Delta_{1n}\, e^{i\mathbf{k}\cdot\mathbf{r}} \quad (n = x, y, z), \tag{25a}
$$

$$
\mathcal{K}_1^2 = -e^{-i\mathbf{k}\cdot\mathbf{r}}\Delta_1^2\, e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{25b}
$$

$$
\mathcal{K}_{2n} = -ie^{-i\mathbf{k}\cdot\mathbf{r}}\Delta_{2n}\, e^{i\mathbf{k}\cdot\mathbf{r}} = \mathcal{K}_{1n}{}^{*} \quad (n = x, y, z), \tag{25c}
$$

$$
\mathcal{K}_2^2 = -e^{-i\mathbf{k}\cdot\mathbf{r}}\Delta_2^2\, e^{i\mathbf{k}\cdot\mathbf{r}} = \mathcal{K}_1^{2*}, \tag{25d}
$$

where the subscripts 1 and 2 on $\mathcal{K}$ indicate the atom type on which the difference is centered. On a diamond/zincblende lattice, using Eq. (8) and the finite differences defined in Eqs. (12a)–(13d), we obtain

$$
H = \begin{pmatrix}
E_{c1} + \mathcal{V}_c & 0 & 0 & 0 & \mathcal{T}_{c1} & iP_{a_1}\mathcal{K}_{1x} & iP_{a_1}\mathcal{K}_{1y} & iP_{a_1}\mathcal{K}_{1z} \\
0 & E_{v1} + \mathcal{V}_v & 0 & 0 & -iP'_{a_1}\mathcal{K}_{1x} & \mathcal{T}_{v1} & iQ_a\mathcal{K}_{1z} & iQ_a\mathcal{K}_{1y} \\
0 & 0 & E_{v1} + \mathcal{V}_v & 0 & -iP'_{a_1}\mathcal{K}_{1y} & iQ_a\mathcal{K}_{1z} & \mathcal{T}_{v1} & iQ_a\mathcal{K}_{1x} \\
0 & 0 & 0 & E_{v1} + \mathcal{V}_v & -iP'_{a_1}\mathcal{K}_{1z} & iQ_a\mathcal{K}_{1y} & iQ_a\mathcal{K}_{1x} & \mathcal{T}_{v1} \\
\mathcal{T}_{c2} & iP_{a_2}\mathcal{K}_{2x} & iP_{a_2}\mathcal{K}_{2y} & iP_{a_2}\mathcal{K}_{2z} & E_{c2} + \mathcal{V}_c & 0 & 0 & 0 \\
-iP'_{a_2}\mathcal{K}_{2x} & \mathcal{T}_{v2} & -iQ_a\mathcal{K}_{2z} & -iQ_a\mathcal{K}_{2y} & 0 & E_{v2} + \mathcal{V}_v & 0 & 0 \\
-iP'_{a_2}\mathcal{K}_{2y} & -iQ_a\mathcal{K}_{2z} & \mathcal{T}_{v2} & -iQ_a\mathcal{K}_{2x} & 0 & 0 & E_{v2} + \mathcal{V}_v & 0 \\
-iP'_{a_2}\mathcal{K}_{2z} & -iQ_a\mathcal{K}_{2y} & -iQ_a\mathcal{K}_{2x} & \mathcal{T}_{v2} & 0 & 0 & 0 & E_{v2} + \mathcal{V}_v
\end{pmatrix},
$$

$$\tag{26}$$

with column labels $S_1 \quad X_1 \quad Y_1 \quad Z_1 \quad S_2 \quad X_2 \quad Y_2 \quad Z_2$.

where the subscripts 1, 2 label the atoms within the unit cell and

$$
\mathcal{T}_{bc} = \left(\frac{1}{2} + F_b\right)\frac{\hbar^2}{m_0}\left(\mathcal{K}_a^2 - \frac{32}{a_{latt}^2}\right), \tag{27a}
$$

$$
\mathcal{V}_b = \left(\frac{1}{2} + F_b\right)\frac{\hbar^2}{m_0}\left(\frac{32}{a_{latt}^2}\right), \tag{27b}
$$

where b is the band index (v or c for valence or conduction), c indicates the type of atom (1 or 2), and the momentum matrix elements are given by Eqs. (20c) and (21).

If the crystal lacks inversion symmetry, then $P_{a_1} \neq P'_{a_2}$, $P'_{a_1} \neq P_{a_2}$, and H is not Hermitian. The problem may be remedied by starting with a set of inversion symmetric Bloch functions, in which case $P_{a_1} = P_{a_2} = P'_{a_1} = P'_{a_2}$. The Hamiltonian can still be inversion non-symmetric due to the different potentials on the anion and cation, giving rise to

inversion non-symmetric zone-center envelope functions. An alternative approach is to modify the differencing scheme using the generalized Voronoi cells described at the end of Section IV. The cell size can be adjusted so as to make $P_{a_2} = P'_{a_1}$ and $P'_{a_2} = P_{a_1}$, restoring Hermiticity while maintaining $P_{a_1} \neq P'_{a_1}$ and $P_{a_2} \neq P'_{a_2}$. Deforming the cells will also change the values of the Q_{a_i}, but since $Q_{a_1} = -Q_{a_2}$ the

form of H will not be affected. In a model with more bands, modifying Ω to make H Hermitian would appear to be limited to tuning the P_{a_i}s for just one pair of bands. One could use different Ωs for different bands, or simply set the P_{a_i}s on an *ad hoc* basis.

We may compare the atomistic four-band $k \cdot p$ model of Eq. (26) with the four-band tight-binding Hamiltonian[37]

$$
H_{tb} = \begin{array}{c} \\ \end{array} \begin{array}{cccccccc} S_1 & X_1 & Y_1 & Z_1 & S_2 & X_2 & Y_2 & Z_2 \end{array}
$$

$$
H_{tb} = \begin{pmatrix}
E_{s1} & 0 & 0 & 0 & V_{ss}g_1 & V_{sp}g_2 & V_{sp}g_3 & V_{sp}g_4 \\
0 & E_{p1} & 0 & 0 & -V_{sp}g_2 & V_{xx}g_1 & V_{xy}g_4 & V_{xy}g_3 \\
0 & 0 & E_{p1} & 0 & -V_{sp}g_3 & V_{xy}g_4 & V_{xx}g_1 & V_{xy}g_2 \\
0 & 0 & 0 & E_{p1} & -V_{sp}g_4 & V_{xy}g_3 & V_{xy}g_2 & V_{xx}g_1 \\
V_{ss}g_1^* & -V_{sp}g_2^* & -V_{sp}g_3^* & -V_{sp}g_4^* & E_{s2} & 0 & 0 & 0 \\
V_{sp}g_2^* & V_{xx}g_1^* & V_{xy}g_4^* & V_{xy}g_3^* & 0 & E_{p2} & 0 & 0 \\
V_{sp}g_3^* & V_{xy}g_4^* & V_{xx}g_1^* & V_{xy}g_2^* & 0 & 0 & E_{p2} & 0 \\
V_{sp}g_4^* & V_{xy}g_3^* & V_{xy}g_2^* & V_{xx}g_1^* & 0 & 0 & 0 & E_{p2}
\end{pmatrix},
\tag{28}
$$

where the gs are the standard tight-binding functions, which are related to the $\mathcal{K}$s by

$$
\begin{aligned}
g_1 &= \frac{1}{4}[\exp(i\mathbf{d}_1 \cdot \mathbf{k}) + \exp(i\mathbf{d}_2 \cdot \mathbf{k}) + \exp(i\mathbf{d}_3 \cdot \mathbf{k}) \\
&\quad + \exp(i\mathbf{d}_4 \cdot \mathbf{k})] = 1 - \frac{a_{latt}^2}{32}\mathcal{K}^2, \\
g_2 &= \frac{1}{4}[\exp(i\mathbf{d}_1 \cdot \mathbf{k}) + \exp(i\mathbf{d}_2 \cdot \mathbf{k}) - \exp(i\mathbf{d}_3 \cdot \mathbf{k}) \\
&\quad - \exp(i\mathbf{d}_4 \cdot \mathbf{k})] = \frac{ia_{latt}}{4}\mathcal{K}_x, \\
g_3 &= \frac{1}{4}[\exp(i\mathbf{d}_1 \cdot \mathbf{k}) - \exp(i\mathbf{d}_2 \cdot \mathbf{k}) + \exp(i\mathbf{d}_3 \cdot \mathbf{k}) \\
&\quad - \exp(i\mathbf{d}_4 \cdot \mathbf{k})] = \frac{ia_{latt}}{4}\mathcal{K}_y, \\
g_4 &= \frac{1}{4}[\exp(i\mathbf{d}_1 \cdot \mathbf{k}) - \exp(i\mathbf{d}_2 \cdot \mathbf{k}) - \exp(i\mathbf{d}_3 \cdot \mathbf{k}) \\
&\quad + \exp(i\mathbf{d}_4 \cdot \mathbf{k})] = \frac{ia_{latt}}{4}\mathcal{K}_z.
\end{aligned}
\tag{29}
$$

Equating the matrix elements in Eqs. (26) and (28), the tight-binding and $k \cdot p$ parameters are related by

$$
V_{ss} = -\frac{\hbar^2}{m_0} \frac{32}{a_{latt}^2} \left(\frac{1}{2} + F_c \right),
\tag{30a}
$$

$$
V_{xx} = -\frac{\hbar^2}{m_0} \frac{32}{a_{latt}^2} \left(\frac{1}{2} + F_v \right),
\tag{30b}
$$

$$
V_{s_1 p_2} = 4P_{a_1}/a_{latt},
\tag{30c}
$$

$$
V_{s_2 p_1} = 4P_{a_2}/a_{latt},
\tag{30d}
$$

$$
V_{xy} = 4Q_a/a_{latt},
\tag{30e}
$$

$$
E_{s1} = E_{c1} + \frac{\hbar^2}{m_0} \frac{32}{a_{latt}^2} \left(\frac{1}{2} + F_c \right),
\tag{30f}
$$

$$
E_{p1} = E_{v1} + \frac{\hbar^2}{m_0} \frac{32}{a_{latt}^2} \left(\frac{1}{2} + F_v \right),
\tag{30g}
$$

$$
E_{s2} = E_{c2} + \frac{\hbar^2}{m_0} \frac{32}{a_{latt}^2} \left(\frac{1}{2} + F_c \right),
\tag{30h}
$$

$$
E_{p2} = E_{v2} + \frac{\hbar^2}{m_0} \frac{32}{a_{latt}^2} \left(\frac{1}{2} + F_v \right).
\tag{30i}
$$

We thus find that the atomistic limit of the $k \cdot p$ model is equivalent to a tight-binding model. In order to make this one-to-one correspondence, it was necessary to include spherically symmetric remote band contributions to both the conduction and valence bands and to have different momentum matrix elements on the two atoms (at least in the inversion non-symmetric case). Our inclusion of only limited remote band contributions to the valence band gives the Luttinger model in the spherical approximation with $\bar{\gamma} = \frac{1}{5}(2\gamma_2 + 3\gamma_3)$.

The atomistic limit always adds at least one new parameter, Q_a. For inversion symmetric crystals, the Hamiltonian matrix elements are the same on both atoms, so there are as many parameters as in the original $k \cdot p$ model, plus the additional parameter Q_a. For inversion non-symmetric crystals, the Hamiltonian matrix elements are different on each atom, so the number of diagonal matrix elements is doubled, plus the additional Q_a. In Sec. VIII, we will take a hybrid approach in which the Hamiltonian is not inversion symmetric, but the Bloch basis is chosen to be symmetric. In that case, the matrix elements of $\hat{H}_0$ (cf. Eq. (1a)) will be different on different atoms, but the momentum matrix elements will be the same on each atom since they depend on derivatives of the inversion symmetric Bloch functions. Therefore, the number of diagonal parameters will be doubled, the number of momentum matrix elements will remain the same, and Q_a will be added.

An important feature of the atomistic limit is that the envelope varies within a unit cell even at zone center, and thus modifies the effective Bloch functions. In Eq. (26), we see that Δ^2 couples the two atoms even at $\mathbf{k} = 0$ via the off-diagonal matrix elements $\mathcal{T}_{c1}$, $\mathcal{T}_{v1}$, $\mathcal{T}_{c2}$, and $\mathcal{T}_{v2}$. This results in a doubling of the number of bands over the continuum model, with the additional bands being shifted by an energy on the order of $\hbar^2/m_0 a_{latt}^2$. Since the atomistic model includes two grid sites per unit cell, the envelope functions include wave vectors outside the first Brillouin zone. These states may be interpreted as approximate Bloch functions with different symmetry from the zone-center Bloch functions of the theory. For example, when multiplied by an envelope that changes sign from site to site the anti-bonding S-like Bloch function of the conduction band becomes similar to the bonding S-like state. Of course such a "fake" Bloch function is not the true zone-center Bloch function, but an approximation. This simple model illustrates the basic features of the atomistic limit, but in order to develop more realistic models, we need to examine the inversion non-symmetric case more closely and study the relationship to heterojunctions.

VI. NON HERMITICITY

As we saw in Sec. V, straight-forward application of the atomistic limit to an inversion non-symmetric crystal gives a non-Hermitian Hamiltonian, since the momentum matrix elements are different on different atoms. This problem generally arises when a finite difference operator is multiplied by a spatially varying coefficient, and also occurs at a heterojunction in continuum $k \cdot p$ theory.[13,25,26,38] In this section and Section VII, we will examine this issue in a general framework that is applicable to both continuum $k \cdot p$ models that have been put on a grid and our atomistic model. Consider a Hamiltonian containing a Hermitian differential operator $\mathcal{D}$ which is approximated by a difference operator $\mathbb{D}$ consisting of a weighted sum of the function values on nearby grid sites

$$\mathcal{D}f(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{R}} \approx \mathbb{D}f(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{R}} = \sum_{\mathbf{R'}} d_{\mathbf{RR'}} f_{\mathbf{R'}}, \qquad (31)$$

where $d_{\mathbf{R},\mathbf{R'}}$ are coefficients defining the difference operator, most of which are zero except when $\mathbf{R'}$ and $\mathbf{R}$ are close to each other. Since $\mathbb{D}$ is Hermitian, $\langle g|\mathbb{D}f\rangle = \langle \mathbb{D}g|f\rangle$ and the corresponding difference operator must satisfy

$$\sum_{\mathbf{R},\mathbf{R'}} g_{\mathbf{R}}^* d_{\mathbf{R},\mathbf{R'}} f_{\mathbf{R'}} = \sum_{\mathbf{R},\mathbf{R'}} d_{\mathbf{R},\mathbf{R'}}^* g_{\mathbf{R'}}^* f_{\mathbf{R}} = \sum_{\mathbf{R},\mathbf{R'}} d_{\mathbf{R'},\mathbf{R}}^* g_{\mathbf{R}}^* f_{\mathbf{R'}}, \quad (32)$$

where the sums on $\mathbf{R}$ and $\mathbf{R'}$ range over all lattice sites. Therefore, the Hermiticity of $\mathbb{D}$ requires $d_{\mathbf{R},\mathbf{R'}} = d_{\mathbf{R'},\mathbf{R}}^*$. If the continuum Hamiltonian contains a differential operator multiplied by a spatially varying coefficient $c(\mathbf{r})$, then simply multiplying that operator by $c(\mathbf{R})$ gives

$$c(\mathbf{r})\mathcal{D}f(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{R}} \to c(\mathbf{r})\mathbb{D}f(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{R}} = \sum_{\mathbf{R}} c_{\mathbf{R}} d_{\mathbf{R},\mathbf{R'}} f_{\mathbf{R'}}. \quad (33)$$

Since Hermiticity requires $c_{\mathbf{R}} d_{\mathbf{R},\mathbf{R'}} = [c_{\mathbf{R'}} d_{\mathbf{R'},\mathbf{R}}]^*$, any spatial variation in the magnitude of $c(\mathbf{r})$ spoils the Hermiticity. This problem arises in a one-band model with a spatially varying effective mass as well as in multi-band envelope models with spatially varying parameters. This problem will occur in Eq. (8) for a heterostructure, but in the atomistic limit it will arise even for a bulk material if the atoms differ from one another.

A common solution is to symmetrize over the connected sites[13,25,26]

$$c(\mathbf{r})\mathcal{D}f(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{R}} \to \sum_{\mathbf{R'}} \frac{1}{2}[c_{\mathbf{R}} + c_{\mathbf{R'}}] d_{\mathbf{R},\mathbf{R'}} f_{\mathbf{R'}}. \qquad (34)$$

The symmetrization is applied to Δ_x, Δ_y, and Δ_z, which are then used to construct other operators. This resolution of the problem is not as *ad hoc* as it may seem, since the first derivative is most naturally defined on a link between two sites. For example, on a one dimensional grid with spacing ϵ, the difference between two adjacent sites gives an approximation to the derivative on the link connecting them

$$\partial_x f(x)\Big|_{x=x_0+\epsilon/2} \approx \Delta_x f(x)\Big|_{x=x_0+\epsilon/2} = \frac{1}{\epsilon}\left(f(x_0+\epsilon) - f(x_0)\right). \qquad (35)$$

Therefore, the value of a coefficient multiplying the derivative is naturally defined on the link itself and should be interpolated between the two points being differenced, giving

$$c(x)\partial_x f(x)\Big|_{x=x_0+\epsilon/2} \approx \frac{1}{2\epsilon}\left[c(x_0) + c(x_0+\epsilon)\right]\left[f(x_0+\epsilon) - f(x_0)\right]. \qquad (36)$$

For terms containing variable coefficients and second derivatives, one must also be careful about the operator ordering. Neither $c(x)\partial_x^2$ nor $\partial_x^2 c(x)$ can be made self-adjoint, even in the continuum, however a symmetrized operator such as $\partial_x c(x)\partial_x$ can.[38–40] Therefore, we can write

$$\partial_x c(x)\partial_x f(x)\Big|_{x=x_0} \approx \frac{1}{\epsilon}\left[c_x \Delta_x f_x\Big|_{x=x_0+\epsilon/2} - c_x \Delta_x f_x\Big|_{x=x_0-\epsilon/2}\right]$$

$$\approx \frac{1}{2\epsilon^2}\left[(c_{x_0} + c_{x_0+\epsilon})f_{x_0+\epsilon} + (c_{x_0} + c_{x_0-\epsilon})f_{x_0-\epsilon}\right.$$

$$\left. - (c_{x_0+\epsilon} + c_{x_0-\epsilon} + 2c_{x_0})f_{x_0}\right]. \qquad (37)$$

VII. FINITE VOLUME METHOD

The symmetrization procedure described above has an intuitive appeal; however, a more formal approach will give the same result while providing some additional insight into the problem. The root cause of the non-Hermiticity is that in discretizing Eq. (5a) we assumed the envelopes and their derivatives were constants over $\Omega_{\mathbf{R}}$. Instead, we can make use of the finite volume method[41] in which the divergence theorem is used to convert the volume integral over a cell into a surface integral. Discretizing this modified version results in finite differences over two sites that are multiplied by a quantity defined on the link between the sites. This

means that the coefficient is the same (except for a possible sign) when evaluated on either of the sites, thus guaranteeing Hermiticity.

Let us return to the continuum, but using Eq. (6a) rather than Eq. (5a), and consider the $\mathcal{P}_{mn}(\mathbf{r}) \cdot \nabla f_n(\mathbf{r})$ term. When integrated over a volume around the grid site $\mathbf{R}$

$$\int_{\Omega_{\mathbf{R}}} \mathcal{P}_{mn}(\mathbf{r}) \cdot \nabla f_n(\mathbf{r}) \, d^3r = \int_{\partial\Omega_{\mathbf{R}}} f_n(\mathbf{r})\mathcal{P}_{mn}(\mathbf{r}) \cdot d\mathbf{s}$$
$$- \int_{\Omega_{\mathbf{R}}} f_n(\mathbf{r})\nabla \cdot \mathcal{P}_{mn}(\mathbf{r}) \, d^3r, \quad (38)$$

where $\partial\Omega_{\mathbf{R}}$ is the bounding surface of $\Omega_{\mathbf{R}}$. Since $f_n(\mathbf{r})$ is slowly varying, we can replace it with $f_{n\,\mathbf{R}}$ in the second integral on the right to obtain

$$\int_{\Omega_{\mathbf{R}}} \mathcal{P}_{mn}(\mathbf{r}) \cdot \nabla f_n(\mathbf{r}) \, d^3r \approx \int_{\partial\Omega_{\mathbf{R}}} f_n(\mathbf{r})\mathcal{P}_{mn}(\mathbf{r}) \cdot d\mathbf{s}$$
$$- f_{n\,\mathbf{R}}\int_{\partial\Omega_{\mathbf{R}}} \mathcal{P}_{mn}(\mathbf{r}) \cdot d\mathbf{s}. \quad (39)$$

The surface $\partial\Omega_{\mathbf{R}}$ is a polyhedron centered on the point $\mathbf{R}$ with faces $S_{\mathbf{d}}$, each of which is normal to the displacement from $\mathbf{R}$ to the neighboring site at $\mathbf{R} + \mathbf{d}$ (see Fig. 1). Because f is slowly varying, its value on one of the faces $S_{\mathbf{d}}$ is approximately constant, with a value that may be approximated by linearly interpolating between the two sites, $f_n(\mathbf{r})|_{S_{\mathbf{d}}} \approx [f_{n\,\mathbf{R}} + f_{n\,\mathbf{R}+\mathbf{d}}]/2$. This gives

$$\int_{\Omega_{\mathbf{R}}} \mathcal{P}_{mn}(\mathbf{r}) \cdot \nabla f_n(\mathbf{r}) \, d^3r \approx \frac{1}{2}\sum_{\mathbf{d}} [f_{n\,\mathbf{R}+\mathbf{d}} - f_{n\,\mathbf{R}}]\int_{S_{\mathbf{d}}} \mathcal{P}_{mn}(\mathbf{r}) \cdot d\mathbf{s}$$
$$(40a)$$

$$\approx \frac{1}{2}\sum_{\mathbf{d}} P_{\mathbf{R}\mathbf{d}} \, \Delta_{\mathbf{d}} f_{n\,\mathbf{R}}, \quad (40b)$$

where

$$P_{mn\,\mathbf{R}\mathbf{d}} = \int_{S_{\mathbf{d}}} \mathbf{d} \cdot \mathcal{P}_{mn}(\mathbf{r}) \, ds, \quad (40c)$$

and $\Delta_{\mathbf{d}}$ is the forward difference in the $\mathbf{d}$ direction defined by $\Delta_{\mathbf{d}}f_{n\,\mathbf{R}} = [f_{n\,\mathbf{R}+\mathbf{d}} - f_{n\,\mathbf{R}}]/|\mathbf{d}|$. The discretized $k \cdot p$ term now consists of a link connecting sites multiplied by a coefficient defined on the link, which is therefore Hermitian. When approximated with naive finite differences, $\nabla \cdot a_{mn}(\mathbf{r})\nabla f_n(\mathbf{r})$ also becomes non-Hermitian. Applying the same methods as above gives

$$\int_{\Omega_{\mathbf{R}}} \nabla \cdot a_{mn}(\mathbf{r})\nabla f_n(\mathbf{r}) \, d^3r \approx \sum_{\mathbf{d}} A_{mn\,\mathbf{R}\mathbf{d}}\Delta_{\mathbf{d}} f_{n\,\mathbf{R}}, \quad (41a)$$

$$A_{mn\,\mathbf{R}\mathbf{d}} = \int_{S_{\mathbf{d}}} a_{mn}(\mathbf{r})ds. \quad (41b)$$

Using the finite volume method, a derivative term becomes a sum of differences between sites multiplied by a coefficient that depends on an integral over the surface separating the cells centered on the two sites. Therefore, the matrix element connecting two sites will be the same whether evaluated at $\mathbf{R}$ or $\mathbf{R} + \mathbf{d}$, making H Hermitian. The quantities given by Eqs. (41b) and (40c) do not need to be explicitly computed, and we may simply use the naive differencing formulas with coefficients empirically fit to bulk properties. The finite volume method provides the justification for the coefficient being determined by the two sites being connected.

We may see explicitly how the finite volume method works by considering the Hamiltonian matrix elements involving the S and X states on two nearest neighbor sites at $\mathbf{R}_1$ and $\mathbf{R}_2$. Using the differences given by Eqs. (12a) and (13a) in Eq. (8), we would have

$$H = \frac{1}{a_{latt}} \frac{\hbar}{im_0} \begin{array}{c} \quad\quad S_1 \quad\quad\quad X_1 \quad\quad\quad S_2 \quad\quad\quad X_2 \\ \begin{pmatrix} 0 & 0 & 0 & \langle S|\hat{p}_x|X\rangle_{\Omega_1} \\ 0 & 0 & \langle X|\hat{p}_x|S\rangle_{\Omega_1} & 0 \\ 0 & -\langle S|\hat{p}_x|X\rangle_{\Omega_2} & 0 & 0 \\ -\langle X|\hat{p}_x|S\rangle_{\Omega_2} & 0 & 0 & 0 \end{pmatrix} \end{array}, \quad (42)$$

which is clearly not Hermitian since all four of the non-zero matrix elements are different. Using Eq. (6a) will partially symmetrize Eq. (42) because $\mathcal{P}_{mn} = -\mathcal{P}_{nm}$, but the matrix elements projected to Ω_1 and Ω_2 will still be different if the basis is inversion non-symmetric. Applying the finite volume method then gives

$$H = \frac{1}{a_{latt}} \begin{array}{c} \quad S_1 \quad\quad X_1 \quad\quad S_2 \quad\quad X_2 \\ \begin{pmatrix} 0 & 0 & 0 & \mathcal{P}_{SX} \\ 0 & 0 & \mathcal{P}_{XS} & 0 \\ 0 & -\mathcal{P}_{SX} & 0 & 0 \\ -\mathcal{P}_{XS} & 0 & 0 & 0 \end{pmatrix} \end{array}, \quad (43)$$

where we have omitted the coordinate indices on $\mathcal{P}$. Since $\mathcal{P}_{mn}$ is real and $\mathcal{P}_{mn} = -\mathcal{P}_{nm}$, H is Hermitian.

There are actually two separate symmetrizations: Eq. (6a) and the finite volume method. The former antisymmetrizes the momentum matrix element with respect to the band index ("index symmetrization") and the latter symmetrizes with respect to the coordinate ("spatial symmetrization"). With an inversion symmetric Bloch basis only the index symmetrization is necessary to obtain a Hermitian Hamiltonian, but in the inversion non-symmetric case the finite volume method provides spatial symmetrization. The finite volume method results in equal momentum

matrix elements, unlike the generalized Voronoi cell approach used in Sec. V which gave two distinct S-X momentum matrix elements for an inversion non-symmetric basis. It is interesting that Eq. (6a) gives a diagonal kinetic term with the same form as the common operator ordering choice in the effective mass approximation, $\nabla \cdot \frac{1}{m(\mathbf{r})} \nabla \psi(\mathbf{r})$.[38–40]

Note that if the Bloch functions centered on $\mathbf{R}$ and $\mathbf{R} + \mathbf{d}$ are different, there will be a discontinuity in the slope of $u_n(\mathbf{r})$ at the interface and the surface integrals in Eqs. (40c) and (41b) will be ill-defined, depending on whether we use $\mathcal{P}_{mn}(\mathbf{r})$ and $a_{mn}(\mathbf{r})$ from the cell around $\mathbf{R}$ or $\mathbf{R} + \mathbf{d}$. Since Bloch functions do not vary greatly among different III–V semiconductors, we may take the surface integral to simply be a parameter depending on the atomic species at $\mathbf{R}$ and $\mathbf{R} + \mathbf{d}$. Moreover, the true u_m at a heterojunction will be a smooth function depending primarily on the type of the dimer, with some small dependence on the nearby atoms. Taking the coefficient to depend only on the two atoms connected is essentially ignoring the influence of nearby atoms on the microscopic $V_0(\mathbf{r})$ and assuming the Bloch function at a dimer is the same as what it would be for that dimer in a bulk binary material. The discontinuity in the Bloch functions at a heterojunction potentially spoils current conservation which requires $\psi(\mathbf{r})$ have continuous first derivatives. A discontinuity in the slope of u_m requires a compensating discontinuity in the slope of the associated envelope. Such an envelope is certainly possible in the continuum, however this cannot be accomplished in a real-space formulation on a grid since such a discontinuity would be over a length scale smaller than the cutoff imposed by the grid. The Burt-Foreman formulation of envelope theory resolves this problem and ensures current conservation, but requires additional momentum matrix elements between wave vectors outside the first Brillouin zone.[42,43]

VIII. EIGHT-BAND MODEL

We now apply the ideas developed in Secs. I–VII to the eight-band Kane model with spin-orbit coupling and perturbative remote band contributions. This model has been used to describe electronic states in bulk materials, impurities, and nanostructures. The Hamiltonian is given by[44–46]

$$
H_8 = \begin{pmatrix}
A & 0 & T^*+V^* & 0 & -\sqrt{3}(T-V) & \sqrt{2}(W-U) & (W-U) & \sqrt{2}(T^*+V^*) \\
0 & A & \sqrt{2}(W-U) & -\sqrt{3}(T^*+V^*) & 0 & T-V & -\sqrt{2}(T-V) & W^*+U \\
(T+V) & \sqrt{2}(W^*-U) & -P+Q & -S^* & R & 0 & \sqrt{\frac{3}{2}}S & -\sqrt{2}Q \\
0 & -\sqrt{3}(T+V) & -S & -P-Q & 0 & R & -\sqrt{2}R & \frac{1}{\sqrt{2}}S \\
-\sqrt{3}(T^*-V^*) & 0 & R^* & 0 & -P-Q & S^* & \frac{1}{\sqrt{2}}S^* & \sqrt{2}R^* \\
\sqrt{2}(W^*-U) & T^*-V^* & 0 & R^* & S & -P+Q & \sqrt{2}Q & \sqrt{\frac{3}{2}}S^* \\
W^*-U & \sqrt{2}(T^*-V^*) & \sqrt{\frac{3}{2}}S^* & -\sqrt{2}R^* & \frac{1}{\sqrt{2}}S & \sqrt{2}Q & Z & 0 \\
\sqrt{2}(T+V) & W+U & -\sqrt{2}Q & \frac{1}{\sqrt{2}}S^* & \sqrt{2}R & \sqrt{\frac{3}{2}}S & 0 & Z
\end{pmatrix},
$$

with basis columns $u^{\Gamma_6}_{-1/2}$, $u^{\Gamma_6}_{1/2}$, $u^{\Gamma_8}_{+1/2}$, $u^{\Gamma_8}_{+3/2}$, $u^{\Gamma_8}_{-3/2}$, $u^{\Gamma_8}_{-1/2}$, $u^{\Gamma_7}_{-1/2}$, $u^{\Gamma_7}_{+1/2}$.

(44)

where

$$
A = E_c + (1+2F)\frac{\hbar^2}{2m_0}k^2, \quad U = \frac{1}{\sqrt{3}}P_0 k_z, \quad V = \frac{1}{\sqrt{6}}P_0(k_x - ik_y), \quad W = \frac{i}{\sqrt{3}}Bk_x k_y,
$$

$$
T = \frac{1}{\sqrt{6}}Bk_z(k_x + ik_y), \quad P = -E_v + \frac{\hbar^2}{2m_0}\gamma_1 k^2, \quad Q = \frac{\hbar^2}{2m_0}\gamma_2 k^2,
$$

$$
R = -\sqrt{3}\frac{\hbar^2}{2m_0}\left(\gamma_2\left(k_x^2 - k_y^2\right) - i2\gamma_3 k_x k_y\right), \quad S = \sqrt{3}\gamma_3 \frac{\hbar^2}{m_0}k_z(k_x - ik_y), \quad Z = E_s - \frac{\hbar^2}{2m_0}\gamma_1 k^2,
$$

(45)

where E_c, E_v, and $E_s = E_v - \Delta$ are the zone-center energies of the conduction, valence, and split-off bands, respectively, F is the remote band contribution to the conduction band effective mass, $iP_0 = \frac{\hbar}{m_0}\langle S|P_x|X\rangle$, and B is the inversion asymmetry parameter due to remote bands, which we take to be zero. The modified Luttinger parameters are given by

$$\gamma_1 = \gamma_1^L - \frac{E_P}{3E_g}, \quad \gamma_2 = \gamma_2^L - \frac{E_P}{6E_g},$$
$$\gamma_3 = \gamma_3^L - \frac{E_P}{6E_g}. \tag{46}$$

Note that we use the definitions from Ref. 44 rather than Ref. 46, since they give the standard relationships between hole masses and Luttinger parameters

$$\left(\frac{m_0}{m_{hh}^*}\right)^{[100]} = \gamma_1^L - 2\gamma_2^L, \quad \left(\frac{m_0}{m_{lh}^*}\right)^{[100]} = \gamma_1^L + 2\gamma_2^L,$$
$$\left(\frac{m_0}{m_{hh}^*}\right)^{[111]} = \gamma_1^L - 2\gamma_3^L, \quad \left(\frac{m_0}{m_{lh}^*}\right)^{[111]} = \gamma_1^L + 2\gamma_3^L. \tag{47}$$

For actual III–V materials, the modified Luttinger parameters in Ref. 46 give effective masses that differ from the relationships given above by about 10%.

To take the atomistic limit of the eight-band model, we proceed as in Sec. V, replacing k's with the appropriate difference operators on the crystal lattice. As in Sec. V, we must include the additional atomistic momentum matrix element $iQ_a = \frac{\hbar}{m_0}\langle X|p_y|Y\rangle_{\Omega_1} = -\frac{\hbar}{m_0}\langle X|p_y|Y\rangle_{\Omega_2}$, where Ω_1 (Ω_2) is the volume around the anion (cation). With this sign convention, a positive Q_a will give a long wavelength 16-band model with $Q > 0$ in agreement with Ref. 33. With the inclusion of spin and transforming to the total angular momentum basis in which Eq. (51) is written, we obtain

$$H_Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\sqrt{3}}Q_a k_- & \frac{1}{\sqrt{3}}Q_a k_z & 0 & -\frac{1}{\sqrt{2}}Q_a k_+ & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}}Q_a k_+ & 0 & 0 & \frac{1}{\sqrt{3}}Q_a k_z & -\sqrt{\frac{2}{3}}Q_a k_z & -\frac{1}{\sqrt{6}}Q_a k_+ \\ 0 & 0 & -\frac{1}{\sqrt{3}}Q_a k_z & 0 & 0 & \frac{1}{\sqrt{3}}Q_a(k_x - ik_y) & \frac{1}{\sqrt{6}}Q_a k_- & -\sqrt{\frac{2}{3}}Q_a k_z \\ 0 & 0 & 0 & -\frac{1}{\sqrt{3}}Q_a k_z & -\frac{1}{\sqrt{3}}Q_a k_+ & 0 & 0 & \frac{1}{\sqrt{2}}Q_a k_- \\ 0 & 0 & \frac{1}{\sqrt{2}}Q_a k_- & \sqrt{\frac{2}{3}}Q_a k_z & -\frac{1}{\sqrt{6}}Q_a k_+ & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{6}}Q_a k_- & \sqrt{\frac{2}{3}}Q_a k_z & -\frac{1}{\sqrt{2}}Q_a k_+ & 0 & 0 \end{pmatrix}, \tag{48}$$

where $k_\pm = (k_x \pm ik_y)$. As in the case of the four-band model, H_Q is the atomistic contribution on the anion site with $-H_Q$ on the cation site.

Rather than using modified Voronoi cells as we did with the four-band model, here we avoid the Hermiticity problem by choosing a basis of inversion symmetric Bloch functions. The eight-band model with $B = 0$ is inversion symmetric, so adopting such a basis is quite natural, and in any case the choice of basis is arbitrary. Although the basis is symmetric, inversion symmetry is broken by $\hat{H}_0$ (cf. Eq. (1a)) which will cause the envelope functions to be inversion non-symmetric. The eight-band model depends on eight parameters (E_c, E_v, E_s, P_0, γ_1, γ_2, γ_3, and F). As discussed in Sec. V, the diagonal energies will be different on each atom, doubling their number to six. The other parameters, which depend on derivatives of the inversion symmetric Bloch functions, will be the same on each atom. With the inclusion of Q_a, the number of parameters is increased from eight to 12 in the atomistic limit.

The remote band contributions in the eight-band model introduce additional ks, and thus additional difference operators. As discussed in Sec. III, the fact that each atom has 4 neighbors means that only k^2, k_x, k_y, and k_z can be constructed using nearest neighbor differences, while $k_x k_y$, $k_x k_z$, $k_y k_z$, k_x^2, k_y^2, and k_z^2 require next nearest neighbor differences. Replacing the ks with difference operators acting on plane waves as in Sec. V, the Hamiltonian becomes a 16×16 matrix of the form

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{12}^\dagger & H_{22} \end{pmatrix}, \tag{49}$$

where the diagonal blocks H_{11} and H_{22} include on-site and next-nearest-neighbor couplings, and the H_{12} block contains nearest neighbor couplings. For $k = 0$, the nearest neighbor couplings from k_x, k_y, and k_z vanish but those from k^2 remain, shifting the zone-center energies from their eight-band values. This can be seen in Eq. (26) where $\mathcal{T}$s are nonzero even for $\mathbf{k} = 0$. For zincblende, the onsite energies are different on the two atoms (E_{c1}, E_{v1}, E_{s1} on the anion and E_{c2}, E_{v2}, E_{s2} on the cation), breaking the inversion symmetry. With two grid sites per unit cell, the number of bands in the original $k \cdot p$ model is doubled in the atomistic limit and each zone-center state of the continuum model splits into two states: one having an envelope with the same sign on each atom ($++$) and another with opposite signs ($+-$). The $++$ eigenvectors correspond to the original eight-band states (Γ_{6c}, Γ_{7v}, Γ_{8v}), and the $+-$ states correspond to the additional states found in the 16-band model (Γ_{6v}, Γ_{8c}, Γ_{7c}) as shown in Fig. 2.

To understand the band doubling, consider a continuum Hamiltonian having a diagonal term $H = E_c + C_c k^2$ where the subscripts indicate the parameters are for the continuum model. The parameters of the atomistic model are determined by fitting to the parameters of the continuum model. Because ∇^2 couples the two atoms within a unit cell, the atomistic Hamiltonian is a non-diagonal 2×2 matrix even at $\mathbf{k} = 0$

$$\begin{pmatrix} E_{a1} + \dfrac{32}{a_{latt}^2}C_a & -\dfrac{32}{a_{latt}^2}C_a \\[3mm] -\dfrac{32}{a_{latt}^2}C_a & E_{a2} + \dfrac{32}{a_{latt}^2}C_a \end{pmatrix}, \qquad (50)$$

where E_{a1}, E_{a2} are the atomistic energies for atom 1 and 2 (anion and cation, respectively, for zincblende), and C_a is the atomistic coefficient corresponding to C_c in the continuum Hamiltonian. The zone-center eigenvalues and eigenvectors are then

$$E = \frac{32}{a_{latt}^2}C_a + \frac{E_{a1} + E_{a2}}{2} \pm \frac{1}{2}\sqrt{(E_{a2} - E_{a1})^2 + 4096C_a^2/a_{latt}^4}, \qquad (51a)$$

$$f = \begin{pmatrix} A/\sqrt{1 + A^2} \\[2mm] 1/\sqrt{1 + A^2} \end{pmatrix}, \qquad (51b)$$

$$A = \frac{a_{latt}^2}{64C_a}(E_{a2} - E_{a1}) \mp \sqrt{\frac{a_{latt}^4}{4096C_a^2}(E_{a2} - E_{a1})^2 + 1}. \qquad (51c)$$

In the inversion symmetric case, $E_{a1} = E_{a2}$ and the zone center energies are given by $E = E_a$ and $E = E_a + 32C_a/a_{latt}^2$. The atomistic parameter E_a is then set equal to the continuum zone center energy, E_c, and the atomistic limit gives rise to an additional band with zone-center energy $32C_a/a_{latt}^2$ higher. If C_a is fixed by fitting to an effective mass, there are no additional parameters to fit. In the inversion non-symmetric case, the zone center energies of the atomistic model are functions of three parameters, E_{a1}, E_{a2}, and C_a, the last of which would be fixed by fitting to an effective mass. We require two conditions to fit E_{a1} and E_{a2}. The simplest approach is to set Eq. (51a) to E_c and the energy of one additional band. Alternatively, one could fit to E_c, and another condition such as the ratio of envelope functions on the two atoms. Since empirical data for energies of excited bands are more readily available than Bloch functions, we will determine E_{a1} and E_{a2} by fitting to the two band energies. Our fitting procedure is by no means unique, and additional data may make some other method preferable.

We take as our target energies those of the 16-band model, E_{6v}, E_{7v}, E_{8v}, E_{6c}, E_{7c}, E_{8c} (see Fig. 2). These have been measured for some materials[47] and calculated for others.[2] Setting the zone-center eigenvalues of the atomistic Hamiltonian equal to the target energies gives the atomistic on-site energies

$$
\begin{aligned}
E_{c1} &= \frac{1}{2}(E_{6c} + E_{6v}) - 16(1 + 2F_a)\hbar^2/m_0 a_{latt}^2 - \frac{1}{2}\sqrt{(E_{6c} - E_{6v})^2 - 1024(1 + 2F_a)^2\hbar^4/m_0^2 a_{latt}^4}, \\[2mm]
E_{c2} &= \frac{1}{2}(E_{6c} + E_{6v}) - 16(1 + 2F_a)\hbar^2/m_0 a_{latt}^2 + \frac{1}{2}\sqrt{(E_{6c} - E_{6v})^2 - 1024(1 + 2F_a)^2\hbar^4/m_0^2 a_{latt}^4}, \\[2mm]
E_{s1} &= \frac{1}{2}(E_{7c} + E_{7v}) + 16\gamma_{a1}\hbar^2/m_0 a_{latt}^2 - \frac{1}{2}\sqrt{(E_{7c} - E_{7v})^2 - 1024\gamma_{a1}^2\hbar^4/m_0^2 a_{latt}^4}, \\[2mm]
E_{s2} &= \frac{1}{2}(E_{7c} + E_{7v}) + 16\gamma_{a1}\hbar^2/m_0 a_{latt}^2 + \frac{1}{2}\sqrt{(E_{7c} - E_{7v})^2 - 1024\gamma_{a1}^2\hbar^4/m_0^2 a_{latt}^4}, \\[2mm]
E_{v1} &= \frac{1}{2}(E_{8c} + E_{8v}) + 16\gamma_{a1}\hbar^2/m_0 a_{latt}^2 - \frac{1}{2}\sqrt{(E_{8c} - E_{8v})^2 - 1024\gamma_{a1}^2\hbar^4/m_0^2 a_{latt}^4}, \\[2mm]
E_{v2} &= \frac{1}{2}(E_{8c} + E_{8v}) + 16\gamma_{a1}\hbar^2/m_0 a_{latt}^2 + \frac{1}{2}\sqrt{(E_{8c} - E_{8v})^2 - 1024\gamma_{a1}^2\hbar^4/m_0^2 a_{latt}^4},
\end{aligned}
\qquad (52)
$$

where E_{c1} is the on-site conduction band energy on the atom at the origin (anion for zincblende), E_{c2} is the on-site energy for the atom at $(a_{latt}/4, a_{latt}/4, a_{latt}/4)$ (cation for zincblende), and likewise for the valence band (subscripts $v1$ and $v2$) and the spin-orbit band (subscripts $s1$ and $s2$). The a subscripts on F_a and γ_{a1} indicate they are the atomistic versions of the continuum $k \cdot p$ parameters F and γ_1. The corresponding envelope functions are obtained from Eq. (51b), giving

$$
\begin{aligned}
A_{6v} &= ((E_{c2} - E_{c1})m_0 a_{latt}^2 + \sqrt{1024(1 + 2F_a)^2\hbar^4 + (E_{c2} - E_{c1})^2 m_0^2 a_{latt}^4})/32(1 + 2F_a)\hbar^2, \\[2mm]
A_{6c} &= ((E_{c2} - E_{c1})m_0 a_{latt}^2 - \sqrt{1024(1 + 2F_a)^2\hbar^4 + (E_{c2} - E_{c1})^2 m_0^2 a_{latt}^4})/32(1 + 2F_a)\hbar^2, \\[2mm]
A_{7v} &= -((E_{s2} - E_{s1})m_0 a_{latt}^2 + \sqrt{1024\gamma_{a1}^2\hbar^4 + (E_{s2} - E_{s1})^2 m_0^2 a_{latt}^4})/32\gamma_{a1}\hbar^2, \\[2mm]
A_{7c} &= -((E_{s2} - E_{s1})m_0 a_{latt}^2 - \sqrt{1024\gamma_{a1}^2\hbar^4 + (E_{s2} - E_{s1})^2 m_0^2 a_{latt}^4})/32\gamma_{a1}\hbar^2, \\[2mm]
A_{8v} &= -((E_{v2} - E_{v1})m_0 a_{latt}^2 + \sqrt{1024\gamma_{a1}^2\hbar^4 + (E_{v2} - E_{v1})^2 m_0^2 a_{latt}^4})/32\gamma_{a1}\hbar^2, \\[2mm]
A_{8c} &= -((E_{v2} - E_{v1})m_0 a_{latt}^2 - \sqrt{1024\gamma_{a1}^2\hbar^4 + (E_{v2} - E_{v1})^2 m_0^2 a_{latt}^4})/32\gamma_{a1}\hbar^2.
\end{aligned}
\qquad (53)
$$
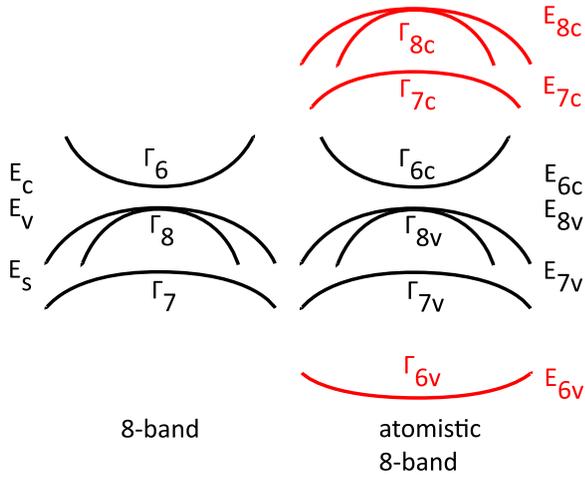
FIG. 2. In the atomistic limit, the number of bands is doubled due to there being two atoms per unit cell. For each band of the original continuum $k \cdot p$ model, there will be one state with an envelope that has the same sign on both atoms $(++)$ and another with an envelope having opposites signs $(+-)$ on the two atoms. The $++$ solutions are taken to be the states of the original continuum $k \cdot p$ model, and the $+-$ solutions, which are shifted in energy by $\approx \mathcal{O}(\hbar^2/m_0 a_{latt}^2)$, are taken to be excited bands. The energy of an excited band also depends on the remote band contribution to the effective mass (i.e., the coefficient of ∇^2 in the atomistic Hamiltonian). The remote contribution to the valence band is multiplied by $\gamma_{a1} > 0$, shifting the Γ_{8c} and Γ_{7c} states to higher energies. Because the remote contribution to the conduction band mass is negative, the Γ_{6v} band is displaced downward.

The states with energies E_{6c}, E_{7v}, E_{8v} have envelopes of the form $++$, while those with energies E_{6v}, E_{8c}, E_{7c} have $+-$ envelopes. Note that the envelopes break inversion symmetry and the momentum matrix elements taken over a unit cell must include the sub-unit cell structure of the envelopes.

Having determined the on-site energies, we must now determine the coefficients of the difference operators. In $k \cdot p$ theory, effective masses only require knowing $E(k)$ to second order in k and are therefore computed using second order perturbation theory in k.[34,48,49] The g-factor is also computed this way.[49] For the 16-band continuum $k \cdot p$ model, the results are

$$\frac{m_0}{m^*} = (1 + 2F) + \frac{E_{P_0}}{3}\left(\frac{2}{E_{6c} - E_{8v}} + \frac{1}{E_{6c} - E_{7v}}\right)$$
$$+ \frac{E_{P_1}}{3}\left(\frac{2}{E_{6c} - E_{8c}} + \frac{1}{E_{6c} - E_{7c}}\right), \quad (54a)$$

$$\frac{g^*}{g_0} = 1 + \frac{g_r}{g_0} - \frac{E_{P_0}}{3}\left(\frac{1}{E_{6c} - E_{8v}} - \frac{1}{E_{6c} - E_{7v}}\right)$$
$$- \frac{E_{P_1}}{3}\left(\frac{1}{E_{6c} - E_{8c}} - \frac{1}{E_{6c} - E_{7c}}\right), \quad (54b)$$

$$\gamma_1^L = \gamma_1 - \frac{E_{P_0}}{3(E_{8v} - E_{6c})} - \frac{E_{P_2}}{3(E_{8v} - E_{6v})}$$
$$- \frac{E_Q}{3(E_{8v} - E_{7c})} - \frac{E_Q}{3(E_{8v} - E_{8c})}, \quad (54c)$$

$$\gamma_2^L = \gamma_2 - \frac{E_{P_0}}{6(E_{8v} - E_{6c})} - \frac{E_{P_2}}{6(E_{8v} - E_{6v})} + \frac{E_Q}{6(E_{8v} - E_{7c})}, \quad (54d)$$

$$\gamma_3^L = \gamma_3 - \frac{E_{P_0}}{6(E_{8v} - E_{6c})} - \frac{E_{P_2}}{6(E_{8v} - E_{6v})} - \frac{E_Q}{6(E_{8v} - E_{7c})}, \quad (54e)$$

where g_0 is the bare electron g factor, g_r is a possible remote band contribution, $E_{P_0} = 2|\langle S_c|P_x|X v\rangle|^2/m_0$, $E_{P_1} = 2|\langle S_c|P_x|X_c\rangle|^2/m_0$, $E_{P_2} = 2|\langle S_v|P_x|X_v\rangle|^2/m_0$, and $E_Q = 2|\langle X_c|P_y|Z_v\rangle|^2/m_0$. The quantities on the left hand sides of Eqs. (54a)–(54e) are the target parameters taken from experiment or possibly *ab initio* calculations, while F, E_{P_0}, E_{P_1}, E_Q, γ_1, γ_2, and γ_3 are the model parameters empirically chosen to reproduce the target values. The model parameters are easily determined, since the equations are linear. Since $E_{8v} - E_{6v}$ is large, we may set $E_{P_2} = 0$, giving five equations in five unknowns.

The effective masses and g-factors of the atomistic model may also be computed using perturbation theory for small k. The resulting expressions contain effective matrix elements that depend on the variation of the envelope over the unit cell and the bands which they connect. For example, the momentum matrix element between the Γ_{6c} and Γ_{8v} states is

$$P_{6c\,8v} = -i\frac{\hbar}{m_0}\langle S_{\Gamma_{6c}}|P_x|X_{\Gamma_{8v}}\rangle$$
$$= \frac{1}{\sqrt{1 + A_{6c}^2}}\begin{pmatrix} A_{6c} \\ 1 \end{pmatrix}^T \begin{pmatrix} 0 & iP_a \\ iP_a & 0 \end{pmatrix} \frac{1}{\sqrt{1 + A_{6c}^2}}\begin{pmatrix} A_{8v} \\ 1 \end{pmatrix}$$
$$= P_a(A_{6c} + A_{8v})/\sqrt{(1 + A_{6c}^2)(1 + A_{8v}^2)}, \quad (55)$$

with the other matrix elements ($P_{6c\,8v}$, $P_{6c\,7v}$, $P_{6c\,8c}$, $P_{6c\,7c}$, $P_{6v\,8v}$, $P_{6v\,7v}$, $P_{6v\,8v}$, and $P_{6v\,7v}$) defined similarly. The Q matrix elements take the form

$$Q_{8c\,8v} = -i\frac{\hbar}{m_0}\langle X_{\Gamma_{8c}}|P_y|Z_{\Gamma_{8v}}\rangle = \frac{(A_{8c} - A_{8v})Q_a}{\sqrt{(1 + A_{8c}^2)(1 + A_{8v}^2)}}. \quad (56)$$

Doing second order perturbation theory directly on the atomistic Hamiltonian, we obtain

$$\frac{m_0}{m_c^*} = \frac{2A_{6c}(1 + 2F_a)}{1 + A_{6c}^2} + \frac{1}{3}\left(\frac{2E_{P\,6c\,8v}}{E_{6c} - E_{8v}} + \frac{E_{P\,6c\,7v}}{E_{6c} - E_{7v}}\right)$$
$$+ \frac{1}{3}\left(\frac{2E_{P\,6c\,8c}}{E_{6c} - E_{8c}} + \frac{E_{P\,6c\,7c}}{E_{6c} - E_{7c}}\right), \quad (57a)$$

$$\frac{g^*}{g_0} = 1 + \frac{g_r}{g_0} - \frac{1}{3}\left(\frac{E_{P\,6c\,8v}}{E_{6c} - E_{8v}} - \frac{E_{P\,6c\,7v}}{E_{6c} - E_{7v}}\right)$$
$$- \frac{1}{3}\left(\frac{E_{P\,6c\,8c}}{E_{6c} - E_{8c}} - \frac{E_{P\,6c\,7c}}{E_{6c} - E_{7c}}\right), \quad (57b)$$

$$\gamma_1^L = \frac{2A_{8v}}{1 + A_{8v}^2}\gamma_{a1} - \frac{E_{P\,8v\,6c}}{3(E_{8v} - E_{6c})} - \frac{E_{P\,8v\,6v}}{3(E_{8v} - E_{6v})}$$
$$- \frac{E_{Q\,8v\,7c}}{3(E_{8v} - E_{7c})} - \frac{E_{Q\,8v\,8c}}{3(E_{8v} - E_{8c})} - \frac{E_{Q\,8v\,7v}}{3(E_{8v} - E_{7v})}, \quad (57c)$$

$$\gamma_2^L = \gamma_{a2} - \frac{E_{P\,8v\,6c}}{6(E_{8v} - E_{6c})} - \frac{E_{P\,8v\,6v}}{6(E_{8v} - E_{6v})}$$
$$+ \frac{E_{Q\,8v\,7c}}{6(E_{8v} - E_{7c})} + \frac{E_{Q\,8v\,7v}}{6(E_{8v} - E_{7v})}, \tag{57d}$$

$$\gamma_3^L = \gamma_{a3} + \frac{A_{8v}Q_a a_{latt}}{6(1 + A_{8v}^2)} - \frac{E_{P\,8v\,6c}}{6(E_{8v} - E_{6c})}$$
$$- \frac{E_{P\,8v\,6v}}{3(E_{8v} - E_{6v})} - \frac{E_{Q\,8v\,7c}}{6(E_{8v} - E_{7c})} - \frac{E_{Q\,8v\,7v}}{6(E_{8v} - E_{7v})}, \tag{57e}$$

where the band-specific Kane energies are $E_{Pmn} = 2|\langle S_m|\hat{p}_x| X_n\rangle|^2/m_0 = 2m_0 P_{mn}^2/\hbar^2$ and $E_{Qmn} = 2|\langle X_m|\hat{p}_y|Z_n\rangle|^2/m_0 = 2m_0 Q_{mn}^2/\hbar^2$. The perturbative expression is the same as in continuum $k \cdot p$ except that matrix elements are replaced with the effective matrix elements including variation of the envelope within the unit cell. Unlike the continuum case, the dependence on F_a and γ_{a1} is nonlinear.

IX. PARAMETER FITTING

To determine the atomistic parameters, we adopt a set of target material parameters to which we fit the atomistic parameters (see Table I). These values are known with varying degrees of certainty, with some taken from high precision measurements while others are obtained theoretically. The basic eight-band parameters are taken from Ref. 50. The zone-center energies of the higher lying bands (Γ_{6v}, Γ_{7c}, Γ_{8c}) are not as well known, and we have taken their values from the $k \cdot p$ calculation of Ref. 6, with the exception of GaAs for which we used the experimental values from Ref. 47. The values of the conduction band effective g factor, g^*, and the Dresselhaus spin splitting, γ_c, were also taken from Ref. 6. Some modifications have been made for InAs. The spin orbit coupling has been increased to $\Delta = 0.45$ eV in order to be able to obtain $g^* = -14.9$ without

having $\gamma_c > 100$ eV Å^3. Alternatively, the Δ could be left unchanged and g^* fit using a remote band contribution g_r. The value of γ_2^L for InAs has been reduced from 8.2 to 7.5 in order to avoid bands that cross the gap at large k, which is much less of a liberty than it may seem since the Luttinger parameters for InAs are poorly known.[50]

The atomistic on-site potentials (E_{c1}, E_{c2}, E_{v1}, etc., in Table II) are determined from Eq. (52) using the lattice constants and zone-center energies from Table I. To fit the effective masses, we must determine the values of F_a, P_a, Q_a, γ_{a1}, γ_{a2}, γ_{a3} for which Eqs. (57a)–(57e) match the empirical target values. This is more difficult than the fitting procedure for a continuum $k \cdot p$ model because the effective momentum matrix elements and remote band contributions depend on F_a and γ_{a1}. We do a nonlinear fit on F_a and γ_{a1}. For particular values of F_a and γ_{a1}, P_a is determined by Eq. (57a), Q_a is determined by Eq. (57c), and γ_{a2} and γ_{a3} are determined by Eqs. (57d) and (57e). F_a is then adjusted to make the resulting g^* match the target value. This results in a curve in the $F_a–\gamma_{a1}$ plane from which we pick the point at which the Dresselhaus spin splitting fits the target as well. We determine γ_c by numerically computing the spin-splitting $E(\mathbf{k})$ in the 110 direction. The range of F_a and γ_{a1} which must be numerically searched is reduced by the condition that the amplitudes A_{6v}, ..., A_{8c} must be real and the solutions corresponding to $E(\Gamma_{6c})$, $E(\Gamma_{7v})$, $E(\Gamma_{8v})$ must have signature $++$. These conditions restrict the values to $-a_{latt}^2(E_{6c} - E_{6v})/32 < (2F_a + 1) < 0$ and $-a_{latt}^2(E_{8c} - E_{8v})/32 < \gamma_{a1} < 0$.

The band structures resulting from our numerical fits are shown in Fig. 3. Since the atomistic model is derived from a continuum $k \cdot p$ model that is perturbative in $\mathbf{k}$, our results are accurate for small $\mathbf{k}$ and all materials appear to have a direct gap. Perturbative $k \cdot p$ models eventually break down at large $\mathbf{k}$, which can result in spurious solutions that cross the gap. When working in a plane wave basis, these spurious solutions may be avoided by simply restricting the values of

TABLE I. Target material parameters. These are the physical material parameters which the atomistic model parameters were adjusted to fit. The values of g^*/g_0 and γ_c from Ref. 6 are experimental values when available, and theoretical values when no measurements were available.

Parameter	AlP	GaP	InP	AlAs	GaAs	InAs	AlSb	GaSb	InSb
a_{latt} (Å)[a]	5.4584	5.4417	5.8613	5.6524	5.6416	6.0501	6.1277	6.0817	6.4690
E_{6v} (eV)[b]	−11.21	−12.14	−11.04	−11.73	−12.9[c]	−11.53	−10.62	−11.47	−10.54
E_{7v} (eV)[a]	−0.07	−0.08	−0.108	−0.28	−0.341	−0.45[d]	−0.676	−0.76	−0.81
E_{8v} (eV)	0	0	0	0	0	0	0	0	0
E_{6c} (eV)[a]	3.63	2.886	1.423	3.099	1.519	0.417	2.386	0.812	0.235
E_{7c} (eV)[e]	4.78	4.38	4.78	4.55	4.488[c]	4.858[d]	3.53	3.11	3.18
E_{8c} (eV)[e]	4.82	4.47	4.97	4.70	4.659[c]	4.79[d]	3.77	3.44	3.64
m^*/m_0 (1)[a]	0.22	0.13	0.0795	0.15	0.067	0.026	0.14	0.039	0.0135
g^*/g_0 (1)[e]	1.92	1.9	1.26	1.52	−0.44	−14.9	0.84	−9.2	−51.6
γ_1^L (1)[a]	3.35	4.05	5.08	3.76	6.98	20	5.18	13.4	34.8
γ_2^L (1)[a]	0.714	0.49	1.6	0.82	2.06	7.5[f]	1.19	4.7	15.5
γ_3^L (1)[a]	1.23	1.25	2.1	1.42	2.93	9.2	1.97	6.0	16.5
γ_c (eV Å^3)[e]	2.1	−2.4	−8.4	11.4	25.0	40.5	40.9	185.0	226.0

[a]Reference 50, except where noted for InAs.
[b]Reference 55, except for InAs.
[c]Reference 47.
[d]Modified to fit g^* and γ_c.
[e]Reference 6, except where noted for GaAs and InAs.
[f]Modified to avoid gap-crossing bands at large $\mathbf{k}$.

TABLE II. Empirical atomistic $k \cdot p$ theory parameters fit to the target parameters in Table I.

Parameter	AlP	GaP	InP	AlAs	GaAs	InAs	AlSb	GaSb	InSb
E_{c1} (eV)	1.560633	0.516670	−1.501874	1.387058	0.675926	−0.354316	−1.094007	−0.568950	−11.520260
E_{c2} (eV)	5.236823	4.669935	3.360307	4.483957	2.273318	1.099526	4.568593	1.934999	−0.968967
E_{v1} (eV)	−0.505855	−0.593772	−0.604112	−0.302607	−0.751431	−0.308368	−0.328926	−0.434964	−0.310302
E_{v2} (eV)	0.645350	0.824859	0.809339	0.347832	1.162880	0.353026	0.400082	0.591758	3.264512
E_{s1} (eV)	−0.620739	−0.597985	−0.591704	−0.836975	−1.198921	−0.788711	−1.423545	−1.490988	−1.213870
E_{s2} (eV)	0.650234	0.689072	0.498932	0.452200	1.098370	−0.056631	0.578701	0.557782	2.808081
F_a (1)	−1.378388	−1.376856	−1.308404	−1.450082	−1.435248	−1.390038	−1.401502	−1.411955	−0.312572
γ_{a1} (1)	−0.571909	−0.514787	−0.671323	−0.609904	−0.554424	−0.731854	−0.569585	−0.498025	−0.117695
γ_{a2} (1)	0.118013	−0.733646	−0.582392	−0.292962	−0.607743	0.572230	0.201367	−0.087151	−1.036618
γ_{a3} (1)	0.139473	−0.334737	−0.221699	−0.010489	−0.113528	0.633312	0.260787	0.246206	−0.582445
P_a (eV Å)	9.325446	10.188130	8.866846	10.192099	10.396101	8.791500	9.428844	10.123984	9.703173
Q_a (eV Å)	−7.062907	−5.962138	−5.085142	−6.279963	−6.139187	−10.995692	−7.559866	−7.470334	−4.704146

k, however this cannot be done in a real-space formulation. Spurious gap-crossing states can be eliminated by modifying the basis,[51,52] choosing different material parameters,[53] or altering the differencing scheme to include higher powers of **k** that push the spurious solutions out of the gap.[54] The threat of gap-crossing bands is greater in the atomistic limit due to the larger (*computational*) Brillouin zone associated with the smaller computational grid, providing more space for the bands to turn over and cross the gap. We show energies throughout the entire Brillouin zone to demonstrate that our parameterization does not produce spurious gap-crossing states, and the model is suitable for use in a real-space formulation. Only InAs required modifications to the parameters to suppress spurious solutions, as described above.

X. CONCLUSION

We have demonstrated how to construct an atomistic $k \cdot p$ theory with finite differences on a grid matched to the crystal lattice. Taking the atomistic limit of $k \cdot p$ theory in a straight-forward way results in a non-Hermitian Hamiltonian, which is seen to be related to the well known fact that multiplying a difference operator by a spatially varying coefficient leads to non-Hermiticity. A more careful treatment shows the problem may be remedied by using the finite volume method, starting with an inversion symmetric Bloch basis, or by deforming the computational cells to generalized Voronoi cells. The use of symmetric Bloch functions does not limit one to the symmetric approximation, since the atomistic envelope functions themselves vary within the unit cell even at **k** = 0. The use of inversion symmetric Bloch functions and generalized Voronoi cells solves the Hermiticity problem, but is not applicable to heterojunctions. As a result, these approaches can be used on systems such as bulk materials, bulk materials with impurities or applied potentials, or nanocrystals with a vacuum barrier. The finite volume method can be used in the presence of heterojunctions.

The atomistic limit of a simple four-band $k \cdot p$ model exactly reproduces the four-band tight-binding model, provided we include spherically symmetric remote band contributions for both the conduction and valence band, and the atomistic momentum matrix elements are different on different atoms (for zincblende). In order to have different momentum matrix elements that make the model exactly match the tight-binding model requires the use of generalized Voronoi cells to symmetrize the momentum matrix elements without making them all equal. The atomistic limit of the widely used eight-band model results reproduces effective masses, g-factors, and Dresselhaus spin splittings of III–V materials. The fits are exact for most materials, with the exception of InAs for which it was necessary to increase the spin-orbit coupling. This may be due to an insufficient number of bands in the model, or due to uncertainties in the experimental values.

The particular implementation presented in Sec. VIII is by no means unique, and different atomistic models are possible depending on the choice of Bloch basis (inversion symmetric or not), the differencing scheme, and whether or not remote band contributions are included. In addition, different fitting procedures may be used. For example, the higher lying band energies could be left as free parameters adjusted to fit the band structure to other criteria such as charge asymmetry. We have chosen to exactly fit all zone center energies, the zone center effective masses for the bottom of the conduction and top of the valence bands, as well as conduction g-factors and Dresselhaus spin splittings, since these quantities are the most important for electronic states of impurities and nanostructures.

An interesting property of these models is that the envelope functions have momenta outside the first Brillouin zone, a feature shared with the Burt-Foreman[42,43] approach to dealing with heterojunctions. Since our model is constructed in real space, there is not a clearly defined separation between the wave function components that are associated with Bloch functions and those that are not, while the Burt-Foreman approach has a clear distinction between Bloch and envelope functions in k-space.

As seen from the fit to III–V materials, the atomistic envelope theory can reproduce the effective masses of the bands near the gap. This is in contrast to tight-binding models which can give incorrect effective masses.[56] The four-band model with only spherically symmetric remote band contributions illustrates how a nearest neighbor tight-binding model fails to reproduce the correct cubic band warping of the valence band. In contrast, the atomistic Kane model gives

(a) AlAs

(b) AlP

(c) AlSb

(d) GaAs

(e) GaP

(f) GaSb

(g) InAs

(h) InP

(i) InSb
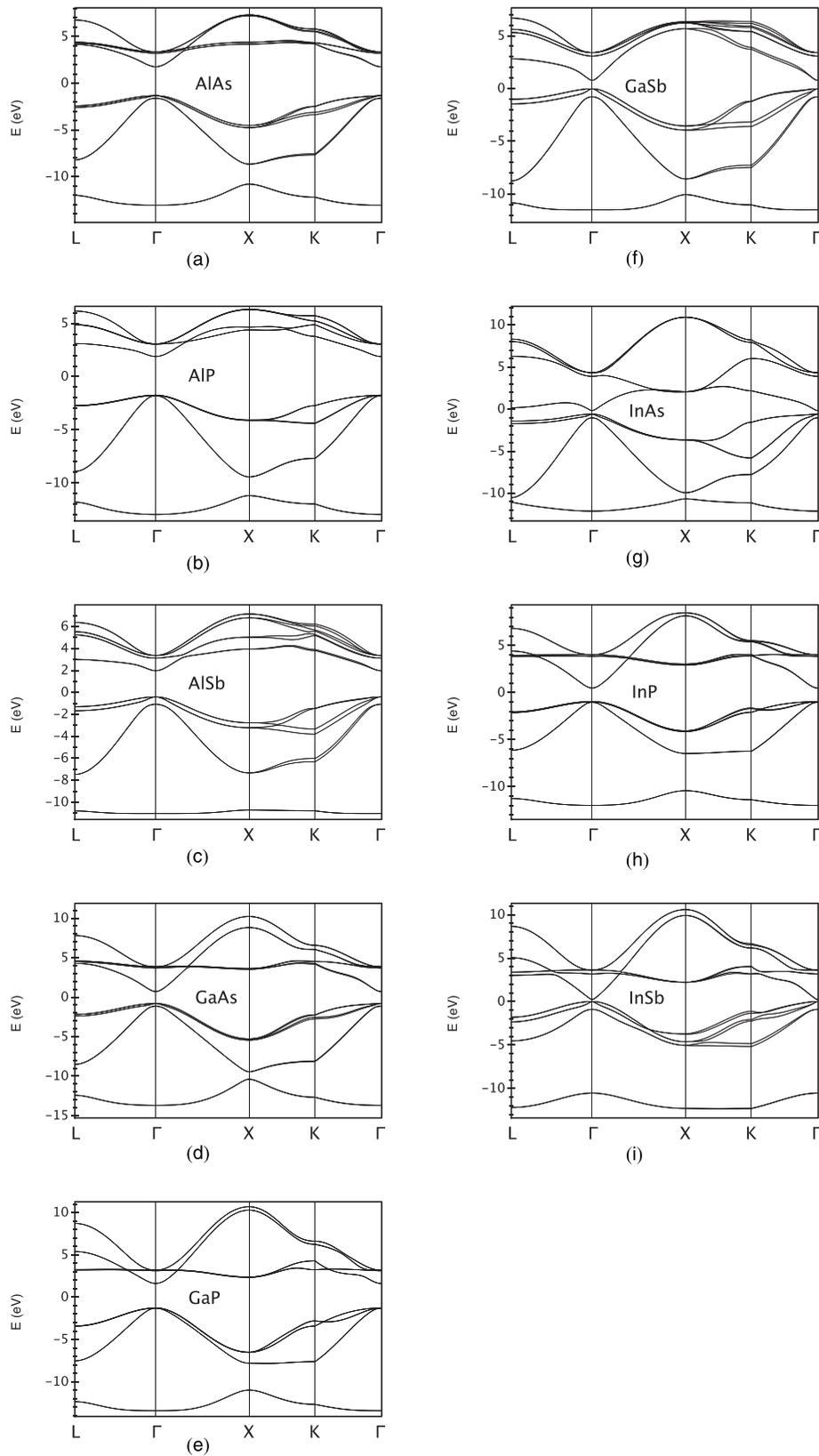
the correct effective masses because it contains next nearest neighbor couplings via the Luttinger parameters.

There are many potential applications of this method to the electronic properties of impurity states, alloys, and poly-types.[57] For sufficiently small nanoparticles, we expect atomistic $k \cdot p$ theory to improve the description of the electronic structure, compared with continuum $k \cdot p$-theory. In particular, nanoparticles with an irregular surface necessitate an atomistic description. It would be interesting to test how well our new method can describe structural defects such as dislocations, twin planes, and stacking defects. Such systems cannot be easily treated in continuum $k \cdot p$-theory.

Atomistic $k \cdot p$ theory will also allow strain effects to be directly modelled in terms of atomic positions, a task which is difficult in both tight-binding and pseudopotential methods. Finally, atomistic $k \cdot p$ theory has the unique feature that it allows the combination of atomistic and continuum models in the same system to facilitate multiscale modeling, since the grid can be highly non-uniform. One could use a rectilinear grid in "large" regions described by a continuum model and an atomistic grid in the regions requiring atomistic precision. The differencing operators in the regions where the rectilinear and atomistic grids meet would be peculiar to the details of the grid used, but would be well defined. Multiscale modeling will dramatically reduce the computing time of atomistic $k \cdot p$-theory compared with other atomistic models, while keeping atomistic accuracy where it is necessary.

ACKNOWLEDGMENTS

APPENDIX: BLOCH BASIS STATES

The Bloch state basis for the eight-band model is

$$u_1 = u_{-1/2}^{\Gamma_6} = |S \downarrow\rangle,$$

$$u_2 = u_{+1/2}^{\Gamma_6} = |S \uparrow\rangle,$$

$$u_3 = u_{+1/2}^{\Gamma_8} = \frac{-i}{\sqrt{6}} |(X + iY) \downarrow\rangle + i\sqrt{\frac{2}{3}} |Z \uparrow\rangle,$$

$$u_4 = u_{+3/2}^{\Gamma_8} = \frac{i}{\sqrt{2}} |(X + iY) \uparrow\rangle,$$

$$u_5 = u_{-3/2}^{\Gamma_8} = \frac{-i}{\sqrt{2}} |(X - iY) \downarrow\rangle,$$

$$u_6 = u_{-1/2}^{\Gamma_8} = \frac{i}{\sqrt{6}} |(X - iY) \uparrow\rangle + i\sqrt{\frac{2}{3}} |Z \downarrow\rangle,$$

$$u_7 = u_{-1/2}^{\Gamma_7} = \frac{-i}{\sqrt{3}} |(X - iY) \uparrow\rangle + \frac{i}{\sqrt{3}} |Z \downarrow\rangle,$$

$$u_8 = u_{+1/2}^{\Gamma_7} = \frac{-i}{\sqrt{3}} |(X + iY) \downarrow\rangle - \frac{i}{\sqrt{3}} |Z \uparrow\rangle,$$

where the ordering of states is the same as for the Hamiltonian in Eq. (51).

[1]M. L. Cohen and T. K. Bergstresser, Phys. Rev. **141**, 789 (1966).

[2]J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1976).

[3]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids **44**, 365 (1983).

[4]J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, Phys. Rev. B **57**, 6493 (1998).

[5]G. Klimeck, R. C. Bowen, T. B. Boykin, and T. A. Cwik, Superlattices Microstruct. **27**, 519 (2000).

[6]J.-M. Jancu, R. Scholz, E. A. de Andrada e Silva, and G. C. La Rocca, Phys. Rev. B **72**, 193201 (2005).

[7]L. L. Y. Voon and M. Willatzen, *The k.p Method: Electronic Properties of Semiconductors* (Springer, 2009).

[8]L. Wang and A. Zunger, Phys. Rev. B **54**, 11417 (1996).

[9]A. Zunger, Phys. Status Solidi B **224**, 727 (2001).

[10]A. J. Williamson and A. Zunger, Phys. Rev. B **59**, 15819 (1999).

[11]T. Saito, J. N. Schulman, and Y. Arakawa, Phys. Rev. B **57**, 13016 (1998).

[12]W. Jaskólski, M. Zieliński, G. W. Bryant, and J. Aizpurua, Phys. Rev. B **74**, 195339 (2006).

[13]M. Grundmann, O. Stier, and D. Bimberg, Phys. Rev. B **52**, 11969 (1995).

[14]M. A. Cusack, P. R. Briddon, and M. Jaros, Phys. Rev. B **54**, R2300 (1996).

[15]H. Jiang and J. Singh, Phys. Rev. B **56**, 4696 (1997).

[16]C. Pryor, Phys. Rev. B **57**, 7190 (1998).

[17]L. Goodwin, A. J. Skinner, and D. G. Pettifor, EPL **9**, 701 (1989).

[18]G. Klimeck, R. C. Bowen, T. B. Boykin, C. Salazar-Lazaro, T. A. Cwik, and A. Stoica, Superlattices Microstruct. **27**, 77 (2000).

[19]J. Kim, L.-W. Wang, and A. Zunger, Phys. Rev. B **57**, R9408 (1998).

[20]M. Cardona and F. H. Pollak, Phys. Rev. **142**, 530 (1966).

[21]N. Fraj, I. Saidi, S. B. Radhia, and K. Boujdaria, Semicond. Sci. Technol. **23**, 085006 (2008).

[22]N. Fraj, I. Saïdi, S. B. Radhia, and K. Boujdaria, J. Appl. Phys. **102**, 053703 (2007).

[23]S. Richard, F. Aniel, and G. Fishman, Phys. Rev. B **72**, 245316 (2005).

[24]I. Saidi, S. Ben Radhia, and K. Boujdaria, J. Appl. Phys. **104**, 023706 (2008).

[25]C. Pryor, Phys. Rev. B **44**, 12912 (1991).

[26]C. Pryor, M.-E. Pistol, and L. Samuelson, Phys. Rev. B **56**, 10404 (1997).

[27]J. K. Cullum and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations* (Birkhäuser, Boston, 1985), Vol. 1–2.

[28]L.-W. Wang and A. Zunger, J. Chem. Phys. **100**, 2394 (1994).

[29]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics Series Vol. 55 (For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964).

[30]B. Fornberg, Math. Comput. **51**, 699 (1988).

[31]W. D. Lakin, Int. J. Numer. Methods Eng. **23**, 209 (1986).

[32]T. Beck, Rev. Mod. Phys. **72**, 1041 (2000).

[33]M. Cardona, N. E. Christensen, and G. Fasol, Phys. Rev. B **38**, 1806 (1988).

[34]P. Pfeffer and W. Zawadzki, Phys. Rev. B **53**, 12813 (1996).

[35]I. D. Mishev, Numer. Methods Partial Differ. Equations **14**, 193 (1998).

[36]A. Telea and J. van Wijk, in *Data Visualization 2001*, edited by D. Ebert, J. Favre, and R. Peikert (Springer, Vienna, 2001), Eurographics, pp. 165–174, ISBN: 978-3-211-83674-3.

[37]D. J. Chadi and M. L. Cohen, Phys. Status Solidi B **68**, 405 (1975).

[38]Q. Zhu and H. Kroemer, Phys. Rev. B **27**, 3519 (1983).

[39]M. S. Birman and M. Z. Solomjak, *Spectral Theory of Self-Adjoint Operators in Hilbert Space* (D. Reidel Publishing Company, Dordrecht, 1986).

[40]G. T. Einevoll, P. C. Hemmer, and J. Thomsen, Phys. Rev. B **42**, 3485 (1990).

[41]P. Knabner and L. Angerman, *Numerical Methods for Elliptic and Parabolic Partial Differential Equations* (Springer, 2003).

[42]M. G. Burt, Semicond. Sci. Technol. **3**, 739 (1988).

[43]B. A. Foreman, Phys. Rev. B **54**, 1909 (1996).

[44]C. Pidgeon and R. Brown, Phys. Rev. **146**, 575 (1966).

[45]E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957).

[46]T. B. Bahder, Phys. Rev. B **41**, 11992 (1990).

[47]D. E. Aspnes and A. A. Studna, Phys. Rev. B **7**, 4605 (1973).

[48]C. Hermann and C. Weisbuch, Phys. Rev. B **15**, 823 (1977).

[49]L. M. Roth, B. Lax, and S. Zwerdling, Phys. Rev. **114**, 90 (1959).

[50]I. Vurgaftman, J. Meyer, and L. Ram-Mohan, J. Appl. Phys. **89**, 5815–5875 (2001).

[51]B. A. Foreman, Phys. Rev. B **56**, R12748 (1997).

[52]B. A. Foreman, Phys. Rev. B **75**, 235331 (2007).

[53]D. Z.-Y. Ting, T. C. McGill, and X. Cartoix, Phys. Rev. B. **68**, 235319 (2003).

[54]M. Holm, M.-E. Pistol, and C. Pryor, J. Appl. Phys. **92**, 932 (2002).

[55]B. D. Malone and M. L. Cohen, J. Phys.: Condens. Matter **25**, 105503 (2013).

[56]T. B. Boykin, Phys. Rev. B **57**, 1620 (1998).

[57]A. De and C. E. Pryor, Phys. Rev. B **81**, 155210 (2010).