Superlattices and Microstructures, Vol. 27, No. 2/3,2000
doi:10.1006/spmi.1999.0797
Available online at http://www.idealibrary.com on IDEAL

Si tight-binding parameters from genetic algorithm fitting

GERHARD KLIMECK, R. CHRIS BOWEN, TIMOTHY B. BOYKIN[‡], CARLOS SALAZAR-LAZARO, THOMAS A. CWIK, ADRIAN STOICA

[†]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, U.S.A.*
[‡]*University of Alabama in Huntsville, Department of Electrical and Computer Engineering, Huntsville, AL 35899, U.S.A.*

(Received 9 November 1999)

Quantum mechanical simulations of carrier transport in Si require an accurate model of the complicated Si bandstructure. Tight-binding models are an attractive method of choice since they bear the full electronic structure symmetry within them and can discretize a realistic device on an atomic scale. However, tight-binding models are not simple to parameterize and characterize. This work addresses two issues: (1) the need for an automated fitting procedure that maps tight-binding orbital interaction energies to physical observables such as effective masses and band edges, and (2) the capabilities and accuracy of the nearest and second-nearest neighbor tight-binding sp3s* models with respect to carrier transport in indirect bandgap materials. A genetic algorithm approach is used to fit orbital interaction energies of these tight-binding models in a nine- and 20-dimensional global optimization problem for Si. A second-nearest neighbor sp3s* parameter set that fits all relevant conduction and valence band properties with a high degree of accuracy is presented. No such global fit was found for the nearest neighbor sp3s* model and two sets, one heavily weighed for electron properties and the other for hole properties, are presented. Bandstructure properties relevant for electron and hole transport in Si derived from these three sets are compared with the seminal Vogl *et al*. [Journal of the Physics and Chemistry of Solids **44**, 365 (1983)] parameters.

Key words: tight binding, sp3s*, silicon, genetic algorithm, effective mass.

1. Introduction

The appeal and problems of tight-binding models. Nanoscaled electronic devices are characterized by material and charge-density variations on the length scale of a few atoms. Tight-binding models can resolve spatial material variations on an atomic scale and they bear the full crystalline and electronic symmetry of semiconductor materials within them. This ability to provide spatial resolution on an atomic scale with complete electronic structure symmetry has led to an increased use of these tight-binding models for the simulation of nanoscaled electronic devices [1–8]. While the tight binding approach is systematically appealing, it bears a big problem in that the basic building constructs for the tight-binding Hamiltonian are not conduction band edges and effective masses that are typically used for electronic device modeling, but orbital interaction energies. The coupling of the discrete orbitals with these interaction energies results in

0749–6036/00/020077 + 12 $35.00/0

the formation of electronic bands. The correct selection of the kind and number of basic orbitals (typically s, p, and d) and their local and nonlocal coupling strength results in the proper representation of material characteristics such as band energies and effective masses. The global bandstructure and effective masses are therefore related to the interaction energies in a nontrivial manner.

Tight-binding parameters for global fits. Given the basic concept of the tight-binding models, the first task was to verify that these models can represent electronic bandstructure properties that can be determined by measurement or other theoretical calculations [9]. The first systematic attempts [10] to evaluate the tight-binding models focussed on the parameter fitting to match global bandstructure data for bands which can be probed by optical measurements. This work has been quite successful in showing that tight-binding models can represent electronic bandstructure with a similar accuracy to other methods. The parameterizations of Vogl *et al*. [10] have provided inroads of the tight-binding models to a variety of fields. In fact, at this moment in time a reference-count engine shows that their paper has been quoted as a reference 716 times [11].

Tight-binding parameters for electronic device modeling. It is important to realize that the tight-binding models do not include all the physics of electronic structure. Only a subset of physical phenomena can be modeled within this framework. The accuracy of these models depends strongly on the choice of orbitals that are included and the parameterization of the orbital interaction energies. It is, for example, understood, but not widely appreciated, that the sp3s* nearest neighbor model pathologically predicts an infinite transverse mass at the X point [5]. Adding more orbitals to the basis or allowing second-nearest neighbor coupling of the sp3s* orbitals eliminates this pathology. However, adding more orbitals or more neighbors to the model adds more orbital interaction energies to the list of parameters that need to be fitted, and the numerical load is increased due to the increase in basis size. With the limitations of the sp3s* model in mind it must be emphasized that early parameterizations [10] provided global band-structure fits. The following two paragraphs serve as a reminder of which bandstructure properties must be properly represented for a quantum mechanical carrier transport simulation.

Masses. The propagation of electrons and holes is dominated by the properties of the lowest conduction and highest three valence bands. Carriers, typically, do not propagate in other valence and conduction bands. However, these other bands shape the symmetries of the conduction and valence bands of interest. The word ‘shape’ is indeed of paramount importance here. The curvature of the bands determines the effective mass, which is the key parameter that determines the propagation of carriers. Most of the electronic device properties such as current, charge densities, and state quantization, scale directly or inversely with the effective mass. Therefore, a bandstructure model used for electronic transport must reproduce the effective mass of the band of interest very well.

Tunneling. An entire class of devices is designed around the physical effect that electrons and holes can penetrate into the bandgap. This quantum mechanical tunneling effect enables the functionality of electronic devices such as tunnel diodes and resonant tunneling diodes. It also determines the current flow through thin CMOS gate dielectrics [8]. The penetration of the electron wavefunction into the bandgap can be described by an exponential decay constant. This decay constant depends directly on the conduction-to-valence-band coupling [6]. Even if a quantum device is designed to be completely electronic in its basic operation (no free holes), such as a typical resonant tunneling diode, the conduction-to-valence-band coupling scales the confinement and tunneling properties exponentially [6].

The need for a fitting algorithm. The success in the quantitative modeling [6, 7] of high-performance resonance tunneling diodes with the Nanoelectronic modeling tool (NEMO) was strongly dependent on the proper parameterization of the sp3s* tight-binding parameters for materials such as InGaAs and InAlAs. To

aid the fitting process, Boykin [12–14] has provided analytical formulas for the effective masses and band edges for nearest and second-nearest sp3s* tight-binding models at the high-symmetry point Γ. This work led to an understanding of which masses can and cannot be fit given the nearest and second-nearest neighbor model. Since the equations are limited to high-symmetry points and since they do not invert into equations for the orbital interaction energies as a function of effective masses and band edges, the fitting remains a tedious task. Another complication arises for materials where the modeling of the X valley parameters is important. The conduction band minimum in the X direction is typically not at the zone edge and no explicit expression for the conduction band minimum and its curvature is known. The parameter optimization is then completely dependent on a numerical technique.

Overview. Section 2 describes our approach to the large dimensional, nonlinear global optimization problem to fit sp3s* tight-binding parameters to experimental observables. Using our resulting parameter sets for Si presented in Section 3 we point out the limitations of the sp3s* nearest and second-nearest neighbor model. Sections 4 and 5 provide an overview of the electronic structure problems that we plan to solve using genetic algorithms (GAs) followed by a summary of this paper, respectively.

2. Optimization procedure

Choice of optimization procedure. Boykin's equations [12–14] for effective masses as a function of interaction energies at high-symmetry points show rich nonlinear behavior. Away from the main symmetry points no explicit effective mass formulas can be derived, but the parameter dependence can also safely be assumed as highly nonlinear. The number of 'free' parameters is large: the nearest and second-nearest neighbor sp3s* tight-binding models for elementary semiconductors such as Si and Ge have nine and 20 orbital interaction parameters, respectively. For compound semiconductors such as GaAs or InP the number of parameters is 15 and 37. Within such a large, nonlinear parameter space a global minimum for a best fit to experimental data is to be found. Given the nonlinearity and size of the problem, it is expected that various local minima exist. This assumption was indeed verified during the course of this research. To address the issue of global minimization we chose a GA approach. Brief and unsuccessful attempts were made to solve our problem using a simulated annealing [15] algorithm.

Genetic algorithm optimization. GA-based optimization employs stochastic methods modeled on the principles of natural selection and the evolution of biological systems [16–18]. They are global, multi-parameter and do not require constraints on continuity of the solution space. They have been introduced in electromagnetic design and modeling over the last 5 years [18–22] in the relatively diverse areas of antenna design, filter design and the design of scattering structures. The work performed on electronic structure modeling and optimization is more recent [23–25]. At JPL work is underway to advance the capabilities for electromagnetic and electronic structure modeling of select microdevice structures. The development of a GA package that is easy and flexible to use forms part of this work.

A GA optimization package consists of a sequence of procedures that lead to an optimized result. The following sequence is common to all GAs with variants at each stage [17]:

- Model parameterization and gene encoding.
- Initialization of population.
- Evaluation of fitness function for population.
- Selection of subset of population.
- Reproduction through crossover and mutation.
- Evaluation of fitness function and convergence check.

This process is shown in diagrammatic form Fig. 1 and is the basis for a general package.

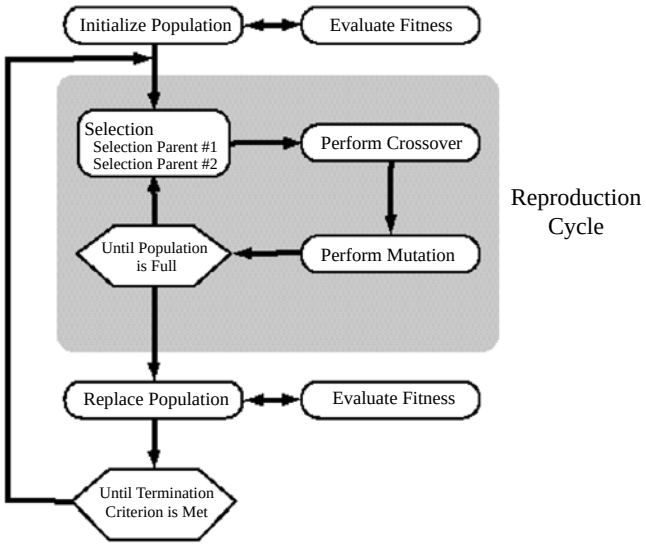

Fig. 1. Genetic algorithm flow diagram (taken from [21]).

Parameterization. One of the most important stages of the optimization process is the sound parameterization of the design and resultant encoding of the parameterization into a chromosome. The parameters need to be chosen from the design space in an effective manner, limited in range to a set that is physical, and encoded in a meaningful way. In the case presented here, for example, the interaction energies do have a physical meaning, in terms of their relative sizes and signs. We have found that for 'blind' optimization, without controlling the allowed range of energies for some of the parameters (such as the spin orbit interaction energies) unphysical results can be obtained. Similarly, how the parameter is encoded into a gene is important. For example, encoding a real parameter into a binary string or using real-valued encoding is itself a tradeoff [6]. In general, efficient convergence of the optimization algorithm will depend heavily upon the initial parameterization and encoding of the design. For the optimization problem presented in this paper a customized real-number encoding scheme was chosen to control the range and mutation size of each parameter more precisely. Each parameter is represented by a double-precision number without further discretization through the typical representation as a binary string.

Choice of a simulation package. A number of GAs and packages exist that fit the structure outlined in Fig. 1. The key requirements for the work presented in this paper are that the package is flexible enough to allow a range of design parameterizations and is able to exploit high-performance computers. First, the different parameterizations and gene encodings, as well as mutation strategies need to be easily available. Secondly, because the parameter space is quite large, a large gene pool needs to be explored. Utilization of massively parallel computers for the fitness evaluation of the independent genes can lead to a dramatic speedup. These points are encapsulated in PGAPack, a parallel genetic algorithm library [26]. This package consists of a set of library routines supplying the user with multiple levels of control over the optimization process. The levels vary from default encodings, with simple initialization of parameters and single-statement execution, to the ability to modify, at a low level, all relevant parameters in the optimization process. User-written routines for evaluation or crossover and mutation can also be inserted if necessary. The package is written using the Message Passing Interface (MPI) for parallel execution on a number of processors. A

master process coordinates the chromosome initialization, selection and reproduction, while slave processes calculate the fitness function, including the execution of the electronic structure code on different processors.

Genetic algorithm parameters. The GA parameters can be classified into two groups: (1) parameter encoding and (2) the GA steering parameters such as population and replacement sizes, crossover and mutation probabilities, stopping and restart conditions. The orbital interaction energies are encoded into double-precision numbers as mentioned above. Each value is drawn randomly from a Gaussian distribution. The mean and standard deviation are entered for each parameter. The Gaussian distribution can be clipped to a hard minimum or maximum. Each value also has a mutation size expressed as a percentage of its current value, attached as an input parameter. For orbital interaction energies that are of the order of single-digit electron volts we typically prescribed a standard deviation of about 0.2–0.4 eV. Smaller orbital interaction energies were typically assigned correspondingly smaller standard deviations. A variety of different population and replacement sizes as well as crossover and mutation probabilities were explored. Satisfactory fitness improvement is typically obtained for populations of 100–200 elements per parameter, a replacement of 10–30% per generation, and mutation and crossover probabilities of 10–50%.

Fitness evaluation. There are two stages to the evaluation of the fitness of a particular parameter set: (1) computation of all the relevant physical quantities, and (2) the assembly of these physical quantities into a single real value that will be used in the optimization for the ranking of this particular set. The details of these two stages for the tight-binding interaction energy optimization problem are given in the following paragraphs.

Bandstructure and effective mass calculation. The NEMO software provides database-like access to materials that can be user defined. User-defined materials are represented in an ASCII character stream that can be parsed by the NEMO material database engine. Once a new material is described by the orbital interaction energies, a variety of material properties can be easily retrieved by simple queries. The database is programmable, where function calls to the standard math library operations ($+$, $-$, sin, cos, min, max, etc.) and a variety of NEMO functions can be included. Typical functions needed for this optimization problem are computation of bandedges, effective masses as a function of band index and electron momentum and the determination of the actual X and L point minima on the X and L axes.

Fitness function. The fitness function provides a ranking algorithm that maps the physics-based data into a *single* number. In the case presented here the fitness function is quite simply the weighted sum of the normalized variances between the targeted and the actual physical values. The GA is driven to minimize this single number. A perfect fit corresponds to a fitness of 0.0. If a 'perfect' fit cannot be obtained the GA is driven towards the smallest positive fitness value. As part of the input to the optimization package a list of physical quantities such as bandedges, momentum minima, and effective masses is provided. Each quantity has an associated weight. The choice of weights influences how individual deviations are prioritized. Setting the weight of a quantity to zero takes that quantity out of the optimization, but enables the monitoring of its values throughout the optimization process. To further restrict the stochastic search we have also included an acceptable minimum and maximum value for the physical quantity. If the computed value falls outside the acceptable range, a fitness value of 10 000 is assigned and all further evaluations of this gene cease.

Required CPU time. The computation of the effective masses and band edges from the tight-binding Hamiltonian takes about 1 second using an SGI Origin 2000 and a HP SPXX-2000. With a population size of 5000, a replacement size of 1000, and an evolution over 1000 generations, this corresponds to about 280 CPU hours. These evolutions were typically run on 16 or 32 CPUs, where communication overhead is still minimal. This corresponds to a wall clock time of about 17.5 and 8.75 hours, respectively.

Table 1: Optimization targets and optimized results for the nearest neighbor (NN) and second-nearest neighbor (2N) model. Two sets of new NN parameters are shown: one optimized for the conduction band (CB) the other for the valence band (VB). The bands are normalized such that $E_v^\Gamma = 0.0$ eV. The effective masses for holes are not listed for the original Vogl *et al*. [10] parameters, since they do not include a spin–orbit interaction.

Property	Target	2N	% dev	NN CB	% dev	NN VB	% dev	Vogl *et al*.	% dev
E_c^Γ	3.350	3.353	−0.098	3.341	0.115	3.346	0.118	3.430	−2.388
$\Delta_{\min}$	0.750	0.758	−1.067	0.597	20.4	0.643	14.26	0.731	2.533
$E_c^{\Delta_{\min}}$	1.130	1.129	0.0496	1.133	−0.257	1.118	1.019	1.171	-3.658
m^*_{Xl}	0.916	0.916	0.030	0.907	1.017	0.531	42.06	0.742	19.07
m^*_{Xt}	0.191	0.191	−0.020	0.297	−56.11	1.054	−453.5	1.620	−750.6
E_c^L	2.050	2.048	0.097	2.031	0.949	2.186	−6.643	2.160	−5.389
$m^*_{lh}[001]$	−0.204	−0.198	3.082	−0.013	93.71	−0.187	8.484		
$m^*_{lh}[011]$	−0.147	−0.146	0.525	−0.012	91.59	−0.154	−4.929		
$m^*_{lh}[111]$	−0.139	−0.139	0.395	−0.010	92.79	−0.117	15.816		
$m^*_{hh}[001]$	−0.275	−0.285	−3.643	−0.030	89.22	−0.348	−26.77		
$m^*_{hh}[011]$	−0.579	−0.581	−0.338	−0.033	94.32	−0.580	−0.174		
$m^*_{hh}[111]$	−0.738	−0.737	0.119	−0.034	95.36	−0.692	6.238		
m^*_{so}	−0.234	−0.237	−1.487	−0.018	92.19	−0.247	−5.381		
Δ_{so}	0.045	0.045	−0.067	0.045	0	0.045	0	0	100

Restarting the genetic algorithm. The GA optimization may get stuck in a local minimum, if a population diversity is not maintained. Restarting the GA may help to re-establish a broad diverse population. We monitor the development of the best parameter set and restart the algorithm if this parameter set does not change for 50 generations. Due to restrictions in the supercomputing queueing system we are also forced to resubmit jobs into the queue that have not reached an optimal value. To preserve the best parameter set available from one run to the next we write out a new input deck every generation. This input deck contains the best parameter set as a starting point of the Gaussian-distributed parameter set. Upon initialization of the population we also ensure that the parameter set fed in from the input is explicitly included in the population.

3. Results: optimized orbital interaction energies

Orbital interaction energies for the nearest and second-nearest neighbor sp3s* model describing Si were optimized with the GA described above. The first two columns of Table 1 show the detailed material properties that were to be properly represented by the tight-binding model. The material properties were taken or derived from Si data published in Ref. [9]. The properties can be characterized into three categories: (1) band edge energies, (2) effective masses in various bands and crystal directions, and (3) crystal momentum ($\Delta_{\min}$) corresponding to the minimum conduction band energy in the X direction. The following sections step through the results of the optimization procedure.

Second-nearest neighbor sp3s.* The third and fourth columns of Table 1 show the material properties that correspond to the optimized second-nearest neighbor ($2N$) model and the relative deviation from the target value in percent. The optimal orbital interaction energies are shown in Table 2. All targeted material properties are represented in the tight-binding model within an accuracy of better than 4%. Twelve out of 15 targets are matched to an accuracy of better than 1%. Such agreement appears to be well within the experimental accuracy of the bandstructure data. The sp3s* second-nearest neighbor model is found to represent the Si material parameters relevant to carrier transport well.

Table 2: Bandstructure model parameters for the nearest-neighbor (NN) and second-nearest neighbor (2NN) model. Two sets of new NN parameters are shown: one optimized for the conduction band (CB) the other for the valence band (VB). All parameters are in units of eV.

Parameter	Vogl *et al.*	NN CB	NN VB	2NN
$E_s(000)$	−4.20000	−3.65866	−3.31789	−4.81341
$E_p(000)$	1.71500	1.67889	1.67862	1.77563
$E_{s^*}(000)$	6.68500	3.87567	8.23164	5.61342
$V_{ss}(\frac{1}{2}\frac{1}{2}\frac{1}{2})$	−8.30000	−7.97142	−9.59895	−8.33255
$V_{xx}(\frac{1}{2}\frac{1}{2}\frac{1}{2})$	1.71500	1.69558	1.69552	1.69916
$V_{xy}(\frac{1}{2}\frac{1}{2}\frac{1}{2})$	4.57500	23.32410	4.77573	5.29091
$V_{sp}(\frac{1}{2}\frac{1}{2}\frac{1}{2})$	5.72920	8.87467	7.14230	5.86140
$V_{s^*p}(\frac{1}{2}\frac{1}{2}\frac{1}{2})$	5.37490	5.41174	7.25052	4.88308
λ_{so}	0.00000	0.04500	0.04500	0.04503
$V_{ss}(110)$				0.01591
$V_{sx}(110)$				0.08002
$V_{sx}(011)$				1.31699
$V_{s^*x}(110)$				−0.00579
$V_{s^*x}(011)$				0.50103
$V_{xx}(110)$				0.00762
$V_{xx}(011)$				−0.10662
$V_{xy}(110)$				0.55067
$V_{xy}(011)$				−2.27784

Nearest Neighbor sp3s.* No such single, well optimized parameter set could be found for the sp3s* nearest neighbor model for Si. With the same optimization weights as used in the second-nearest neighbor model, both the electron and hole masses converged to values significantly different from the experimental target values. Two separate numerical experiments where the weights for electron masses were raised and lowered for the holes, and vice versa, resulted in two different 'optimal' sets. The sets were labeled nearest neighbor (NN) conduction band (CB) and valence band (VB), and the material properties are tabulated in columns 5/6 and 7/8 of Table 1, respectively. The VB parameter set reproduces the hole masses and bands reasonably well. The symmetry of the bands are reproduced properly. This parameter set may be useful for hole simulations using the NN tight-binding sp3s* model.

Even with the focus on electron masses a good agreement with the transverse/light electron mass could not be achieved. The deviation still remained over 50%. Table 2 shows that the CB parameters contain an interaction energy V_{xy} of 23 eV which is completely unphysical. The reason for the usage of tight-binding models in electron transport, which is the proper description of the electron wavefunction symmetry, will be lost. The optimization drives this one parameter to such extreme values due to the pathology of the sp3s* nearest neighbor model. Although the effective masses and bandedges for electrons appear reasonable, we urge for caution when considering the use of the sp3s* NN model to simulate electron transport in Si.

The material parameters that result from the standard Vogl parameters are included in columns 9/10 of Table 1 for reference. These parameters do not include a spin–orbit interaction. The hole bands and masses are, therefore, not properly represented. The conduction band properties agree reasonably well with the experimental data, considering the limits of the model.

Graphical comparisons. Figure 2 depicts the conduction band minimum along the [100] direction from the Γ to X point for the four parameter sets discussed above. All four curves 'look' reasonable in comparison

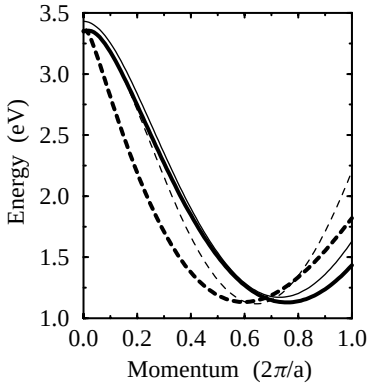

Fig. 2. Conduction bands computed in the nearest neighbor (NN) and second-nearest neighbor (2NN) models as a function of electron momentum in the [100] direction. The NN model is computed with three different parameter sets: electron optimized, hole optimized, and original Vogl *et al*. [10] parameters.

to each other. Indeed, Table 2 shows that the heavy electron masses all agree within 20% to the experimental value. Note, however, that a graph like this does not show any of the transverse momentum dependence that results in the material property of the light-electron mass.

The capability of modeling the light-electron mass in the nearest neighbor and second-nearest neighbor tight-binding model is the theme of Fig. 3. The columns of Fig. 3 are grouped by the parameter sets: (left column) second-nearest neighbor (2N), (center column) nearest neighbor electron optimization (NN CB), and (right column) nearest neighbor hole optimization (NN VB). The left column will be used as a reference since it provides the best agreement with experimental data. The upper row of Fig. 3 shows the bands along the [100] and [111] directions. The symmetry and curvatures shown in 3A correspond to the data shown for Si in Ref. [9]. Figure 3B which stems from the electron optimized parameters shows the lowest conduction band to be similar to 3A. The upper conduction bands have been pushed to lower energies at the Γ point, providing more curvature to the lowest conduction band. In fact, the minimum conduction band momentum ($\Delta_{\min}$) on the [100] line is pushed in from the target value of 0.75, to 0.597. Note that the hole band symmetry is not at all well represented. The hole masses are also clearly too light as indicated by Table 2. Figure 3C shows quite similar behavior to Fig. 3A, in terms of the hole bands and the lowest conduction band, however, the upper conduction bands are quite modified.

The second row of Fig. 3 shows the contour plots of the lowest conduction band in the (001) plane of the first Brillouin zone. The dark blue colors correspond to the minimum of the scale at about 1.13 eV. The red color corresponds to the maximum resolved energy of about 3.36 eV. The third row shows the lowest conduction band isosurfaces in the first Brillouin zone at an energy of 1.45 eV. Figures 3D–F and 3G–I all show the expected four and six conduction band minima. However, only 3D and G show the correct 'cigar' shape representing a correct light-electron and heavy-electron symmetry. Figures 3F and I show 'pancake'-like conduction band surfaces, rather than the expected 'cigar'-shaped surfaces. In fact, the light electron mass is larger than the heavy electron mass (see also Table 1). The shape of the conduction band optimized nearest neighbor (CB NN) parameter set shows roughly the right symmetry with the correct light/heavy electron ratio.

Figure 3F clearly shows the complete pathology of the sp3s* model with respect to the transverse electron mass. From the coloring it can be clearly seen that at the zone faces the band edges are flat and the mass along these edges is indeed infinite. A finite transverse effective mass can only be obtained by pushing the conduction band minimum away from the zone edge. The deeper the conduction band minimum is pushed towards the Γ point, the smaller the transverse effective mass can be made. This comes with the cost of bringing the upper conduction bands much lower in energy and losing the proper masses for the hole bands. As also

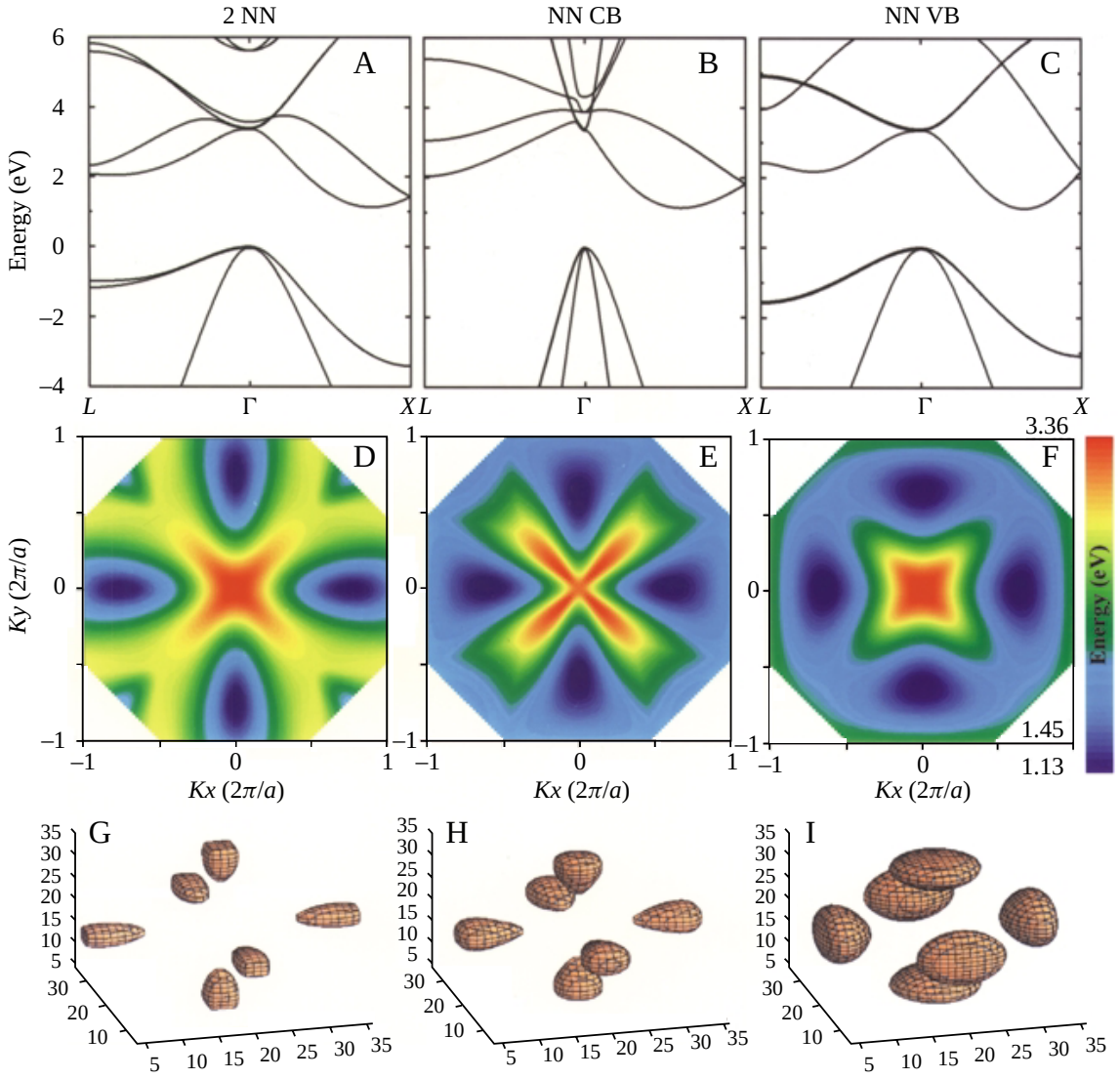

Fig. 3. Bandstructure quantities computed for three different model and parameter combinations: (1) left column, second-nearest neighbor (2N), (2) center column, nearest neighbor optimized for electrons (NN CB), and (3) right column, nearest-neighbor optimized for holes (NN VB). First row: Bands along the [100] and [111] directions. Second row: contour plot of lowest conduction band in the first Brillouin zone in the (001) plane, Third row: Isosurface of the lowest conduction band at an energy of 1.45 eV.

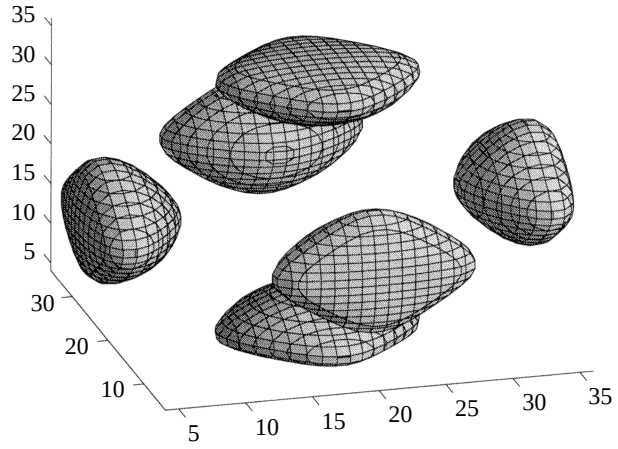

Fig. 4. Lowest conduction band isosurface at an energy of 1.45 eV based on the Vogl *et al*. [10] parameters.

noted above, one of the interaction energies becomes unphysically large and electron and hole wavefunction symmetries are *not* expected to be correct.

For reference, Fig. 4 shows the lowest conduction band isosurface at 1.45 eV in the Brillouin zone computed with the standard Vogl *et al*. [10] parameters. The incorrect shape of the conduction band minimum surface is clearly evident, similar to Fig. 3I.

Which model to use? The second-nearest neighbor sp3s* tight-binding model clearly represents the complex Si bandstructure better than the nearest neighbor model. The superior model comes with the price of added dimensionality in the parameter fitting problem (9 to 20 parameters for elementary semiconductors, 15 to 37 for compound semiconductors) and with the price of increased numerical complexity due to the increase in the size of the basis. Despite its complications the second-nearest neighbor model should be used preferentially over the nearest neighbor model since the latter lacks the basic capabilities to simultaneously model electrons and holes properly.

Use of nearest neighbor model for indirect bandgap materials. Although the transverse mass pathology in the nearest neighbor model is known, some researchers continue to use [27] this model to simulate problems where the proper light-electron mass is important, or where the proper hole-masses are important. We hope that Figs 3 and 4 clearly visualize the limits of the nearest neighbor sp3s* model away from the Γ point. For existing research software, where a rewrite of the underlying algorithms is impossible/impractical, we suggest the usage of the two new parameter sets provided here as a sanity check for parameter dependence in the simulation. For 'pure' hole simulations the specialized nearest neighbor parameter set may provide more physical answers.

Special caution needs to be taken with the electron parameter set, since it does include an unphysically large interaction energy. This parameter set should be taken as a warning that this model should not be used for the modeling of indirect bandgap materials where conduction band minima are in the X or L valley. Forcing the conduction band to a certain curvature off the Γ point will lead to unphysical interaction energies.

4. Future work

The GA-based optimization scheme in its generality and flexibility has enabled us to look at complex parameterizations of tight-binding models for electron transport simulations. The second-nearest neighbor model hides some caveats with respect to three-center interaction integrals and the scaling of these interaction energies with strain becomes quite complicated. We are, therefore, exploring sp3d5 tight-binding models to properly represent the highest valence and lowest conduction bands under the inclusion of strain. The GA approach has enabled us to consider sp3d5 tight binding models seriously, since we can now overcome the fitting 'nightmare'. Preliminary results show that the GA fitting procedure also works well for the sp3s5 models and results will be published elsewhere. A new parameterization of the nearest neighbor sp3s* model for the nine binary compounds consisting of Ga, In, Al and As, Sb, and P is given in [24].

5. Summary

This work has presented three major points needed for the quantum mechanical simulation of carrier transport in Si:

- A genetic algorithm-based optimization scheme for the determination of orbital interaction energies for tight-binding models, given physical observables such as band edges and effective masses.
- Three new parameter sets to represent Si in the nearest and second-nearest sp3s* models
- A clear visualization of the transverse X mass pathology of the nearest neighbor sp3s* model.

We hope that Fig. 3 helps deter the use of the nearest neighbor sp3s* model for the modeling of indirect bandgap materials such as Si or Ge.

Acknowledgements—The work described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration. The supercomputer used in this investigation was provided by funding from the NASA Offices of Earth Science, Aeronautics, and Space Science. Part of the research reported here was performed using HP SPP-2000 operated by the Center for Advanced Computing Research at Caltech; access to this facility was provided by Caltech.

GK would like to highlight and acknowledge the extensive PGAPack work by David Levine [26] including the variety of examples and software manual. His thorough software development enabled our entry into the world of massively parallel GA-based optimization.

References

[1] Y. C. Chang and J. N. Schulman, Phys. Rev. **B25**, 3975 (1982).
[2] J. N. Schulman and Y. C. Chang, Phys. Rev. **B31**, 2056 (1985).
[3] J. N. Schulman, J. Appl. Phys. **60**, 3954 (1986).
[4] D. Z. Y. Ting, E. T. Yu, and T. C. McGill, Phys. Rev. **B45**, 3583 (1992).
[5] T. B. Boykin, J. P. A. van der Wagt, and J. S. Harris, Phys. Rev. **B43**, 4777 (1991).
[6] R. C. Bowen, G. Klimeck, R. K. Lake, W. R. Frensley, and T. Moise, J. Appl. Phys. **81**, 3207 (1997).
[7] G. Klimeck, T. Boykin, R. C. Bowen, R. Lake, D. Blanks, T. S. Moise, Y. C. Kao, and W. R. Frensley, *The 1997 55th Annual Device Research Conference Digest* (IEEE, NJ, 1997) p. 92.
[8] C. Bowen *et al*., *IEDM 1997* (IEEE, New York, 1997) pp. 869–872.
[9] O. Madelung, *Semiconductors-Basic Data* (Springer, Berlin, 1996).
[10] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids **44**, 365 (1983).
[11] Citiation Database, Institute for Scientific Information (`http://www.isinet.com`).
[12] T. B. Boykin, G. Klimeck, R. C. Bowen, and R. K. Lake, Phys. Rev. **B56**, 4102 (1997).

[13] T. B. Boykin, Phys. Rev. **B56**, 9613 (1997).
[14] T. B. Boykin, L. J. Gamble, G. Klimeck, and R. C. Bowen, Phys. Rev. **B59**, 7301 (1999).
[15] Goffe, Ferrier, and Rogers, J. Econometrics **60**, 65 (1994).
[16] J. Holland, *Adaptation in Natural and Artificial Systems: an introductory analysis with applications to biology, control and artificial intelligence* (The University of Michigan Press, Ann Arbor, 1975).
[17] D. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, New York, 1989).
[18] D. S. Weile and E. Michielssen, IEEE Trans. Antennas Propag. **AP-45**, 343 (1997).
[19] E. Michielssen, S. Ranjithan, and R. Mittra, *Optimal Multilayer Filter Design Using Real Coded Genetic Algorithms* (Proceedings of the Institute of Electronic Engineering, ADDRESS, 1992) pp. 413–420.
[20] R. Haupt, IEEE Trans. Antennas Propag. **AP-42**, 939 (1994).
[21] J. Johnson and Y. Rahmat-Samii, IEEE Antennas Propag. Mag. **939**, 7 (1997).
[22] C. Zuffada and T. Cwik, IEEE Trans. Antennas Propag. **46**, 657 (1998).
[23] F. Starrost, S. Bornholdt, C. Solterbeck, and W. Schattke, Phys. Rev. **B53**, 12549 (1996).
[24] G. Klimeck, R. C. Bowen, T. B. Boykin, and T. A. Cwik, Superlatt. Microstruct., to be published, doi:10.1006/spmi.2000.0822.
[25] G. Klimeck, C. Salazar-Lazaro, A. Stoica, and T. Cwik, in *Genetically engineered nanostructure devices: Materials Research Society, Symposium Proceedings*, Materials in Space—Science, Technology and Exploration, edited by A. F. Hepp, J. M. Prahl, T. G. Keith, S. G. Bailey, and J. R. Fowler (Materials Research Society, Warrendale, PA, 1999) Vol. 551, p. 149.
[26] D. Levine, Technical Report No. 95/18, Argonne Nat. Lab. (unpublished) ftp://ftp.mcs.anl.gov/pub/pgapack/.
[27] K. Leung and K. B. Whaley, Physical Review **B56**, 7455 (1997).