Coarse-grained tight-binding models

Tian-Xiang Liu1, Li Mao1, *, Mats-Erik Pistol2, *, and Craig Pryor3, *

1 School of Physics and Technology, Wuhan University, Wuhan 430072, People’s Republic of China
2 NanoLund and Solid State Physics, Lund University, PO Box 118, 221 00 Lund, Sweden
3 Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa 52242, United States of America

E-mail: maoli@whu.edu.cn, mats-erik.pistol@ftf.lth.se and craig-pryor@uiowa.edu

Received 3 October 2021, revised 25 November 2021
Accepted for publication 17 December 2021
Published 5 January 2022

Abstract
Calculating the electronic structure of systems involving very different length scales presents a challenge. Empirical atomistic descriptions such as pseudopotentials or tight-binding models allow one to calculate the effects of atomic placements, but the computational burden increases rapidly with the size of the system, limiting the ability to treat weakly bound extended electronic states. Here we propose a new method to connect atomistic and quasi-continuous models, thus speeding up tight-binding calculations for large systems. We divide a structure into blocks consisting of several unit cells which we diagonalize individually. We then construct a tight-binding Hamiltonian for the full structure using a truncated basis for the blocks, ignoring states having large energy eigenvalues and retaining states with energies close to the band edge energies. A numerical test using a GaAs/AlAs quantum well shows the computation time can be decreased to less than 5% of the full calculation with errors of less than 1%. We give data for the trade-offs between computing time and loss of accuracy. We also tested calculations of the density of states for a GaAs/AlAs quantum well and find a ten times speedup without much loss in accuracy.

Keywords: coarse-grained, tight-binding models, semiconductors

(Some figures may appear in colour only in the online journal)

1. Introduction

There are three major empirical methods for calculating bulk electronic band structures of semiconductors within the single particle approximation: pseudo-potentials [1, 2], \( \vec{k} \cdot \vec{p} \) theory [3, 4] and tight-binding models [5]. When these are extended to complex inhomogeneous systems such as nanostructures or materials with impurities, they differ in how the computational difficulty scales with the system size. All the methods are accurate but involve trade-offs between accuracy, sensitivity to material details, and computational cost for inhomogeneous systems. The empirical pseudo-potential (EPP) method has a large computational burden when applied to large systems, though systems with millions of atoms are feasible. The \( \vec{k} \cdot \vec{p} \) method is a quasi-continuous model that does not explicitly involve the placement of individual atoms except in its atomistic extension which is similar to a tight-binding model [6]. Like the EPP model, the empirical tight-binding (ETB) model is sensitive to the placement of individual atoms [7, 8], but due the local couplings of atomic orbitals it results in a sparse Hamiltonian [9] which is more efficiently diagonalized than the non-sparse Hamiltonian resulting from EPP models in a plane wave basis. A sparse Hamiltonian can be obtained in EPP calculations if done in real space [10].

The most widely used formulation of the ETB method was introduced by Slater and Koster [11] and many band structure studies based on it have been reported [12–16]. The ETB method has been extended to include open boundary conditions [17], electro-magnetic fields [18] and surface passivation [19]. The ETB method is, however, limited in the size of the systems that can be modeled due to the fact that the number of degrees of freedom is proportional to the number of atoms.
To determine the electronic states of a complex inhomogeneous system such as a nanostructure one assumes the Hamiltonian matrix elements are given by the values corresponding to bulk materials. In this way the Hamiltonian matrix may be built up atom by atom and then diagonalized to determine energies and wave functions. Weakly bound states extending over long distances can present difficulties when using an atomistic model. For example, consider an electron bound to a phosphorus donor in a GaAs/AlAs heterostructure in which the heterostructure composition varies over length scales comparable to the effective Bohr radius in GaAs, \( a \approx 10.8 \text{ nm} \). Calculations for such a system would require a computational cell with a volume on the order of \( 10^3 \) to avoid finite size effects, or about \( 5 \times 10^7 \) atoms. An \( spds^* \) model with ten orbitals per atom would thus necessitate a state vector with \( 5 \times 10^8 \) components for this simple problem. These different length scales present a significant computational challenge if both the atomic scale structure near the impurity and the large scale structure are important.

We propose mitigating this problem by coarse-graining a tight-binding model. The system is divided into subunits or blocks of atoms, and then a reduced basis is used within those blocks. The rank of the Hamiltonian is therefore reduced, but in a way that is compatible with the original tight-binding model which may be retained in regions where atomistic precision is required. The resulting Hamiltonian is smaller than the original Hamiltonian, thereby speeding up the calculation and reducing memory requirements with little loss in accuracy. This method bears a resemblance to the blocking transformations used in real-space renormalization [22, 23]. Here, however, there is no continuum limit since the ‘true theory’ is the atomistic tight-binding model.

In section 2 we present the method in detail and section 3 contains numerical results for a quantum well test case. This includes an analysis of the impact of block sizes and energy cutoffs to the accuracy and efficiency of the calculations. A discussion and a summary is given in section 4.

2. Method

In the tight-binding method the wave function is constructed out of atomic orbitals as

\[
\psi(\mathbf{r}) = \sum_{ij} a_{ij} \phi_{ij}(\mathbf{r})
\]

where \( \phi_{ij}(\mathbf{r}) \) is the \( i \)th atomic orbital centered on atom \( j \) and \( a_{ij} \) are the coefficients to be solved for. In the most commonly used approximation the Hamiltonian matrix which acts on \( a_{ij} \) only couples nearest neighbor atoms. These hopping matrix elements have a simple angular dependence due to the neglect of three-center integrals [11]. The atomic orbitals on different atoms are assumed orthogonal, which can be satisfied by using Löwdin’s method [24].

In the ETB model the Hamiltonian matrix elements are determined by fitting to known properties of bulk materials. A primitive unit cell is used to compute the band structure as a function of the Hamiltonian matrix elements which are then adjusted to reproduce the band structure known from experiment or ab initio calculations. To determine the electronic states of a complex inhomogeneous system such as a nanostructure one assumes the Hamiltonian matrix elements are given by the values of the corresponding bulk materials. The effective Schrödinger equation for the \( a_{ij} \) is then

\[
\sum_{kl} H_{ijkl} a_{ij}^{(n)} = E^{(n)} a_{ij}^{(n)}
\]

where the matrix elements \( H_{ijkl} \) are set to the values corresponding to a bulk material having the same atoms as those connected by \( H_{ijkl} \). In this way the Hamiltonian matrix may be built up atom by atom and then diagonalized to determine energies \( E_n \) and wave functions \( a_{ij}^{(n)} \). The form of equation (1) allows wave functions that vary over atomic scales but presents computational difficulties if the wave function

---

Figure 1. Examples of blocks used in our 1D test system. Each block consists of an integer number of unit cells. The top left panel depicts the smallest block used, one unit cell thick. The other blocks are two unit cells and three unit cells in the top right and bottom panel.

Figure 2. Calculations for GaAs quantum wells in AlAs having different thicknesses. Blue dots correspond to the ground state of the electrons and red dots correspond to the ground state of holes in the quantum well. The full Hamiltonian was used for this calculation.
extends over large distances, as illustrated by the example in the introduction.

The details of our method is illustrated by a planar structure such as a quantum well. We begin with a supercell that is one atom thick in the \(x\) and \(y\) directions and has some structure in the \(z\) directions (e.g. a quantum well). There are periodic boundary conditions in all three directions and we include a wave vector \(\vec{k}\). By grouping atoms into blocks as shown in figure 1 the Hamiltonian can be written in the form

\[
\mathcal{H}(\vec{k}) = \begin{bmatrix}
H_0 & V_{01} & 0 & \cdots & 0 & V_{0,n} \\
V_{01}^\dagger & H_1 & V_{12} & \cdots & 0 & 0 \\
0 & V_{12}^\dagger & H_2 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & H_{n-1} & V_{n-1,n} \\
V_{0,n}^\dagger & 0 & 0 & \cdots & V_{n-1,n}^\dagger & H_n
\end{bmatrix}
\]  

(3)

where \(H_i\) is the sub-matrix involving those atoms in the \(i\)th block and \(V_{ij}\) is the sub-matrix coupling the \(i\)th and \(j\)th blocks.

We then apply a unitary transformation to obtain

\[
\tilde{\mathcal{G}}(\vec{k}) = \begin{bmatrix}
\tilde{G}_0 & \tilde{g}_{01} & 0 & \cdots & 0 & \tilde{g}_{0,n} \\
\tilde{g}_{01}^\dagger & \tilde{G}_1 & \tilde{g}_{12} & \cdots & 0 & 0 \\
0 & \tilde{g}_{12}^\dagger & \tilde{G}_2 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \tilde{G}_{n-1} & \tilde{g}_{n-1,n} \\
\tilde{g}_{0,n}^\dagger & 0 & 0 & \cdots & \tilde{g}_{n-1,n}^\dagger & \tilde{G}_n
\end{bmatrix}
\]  

(4)

where \(\tilde{G}_i = S_i^{-1} H_i S_i\) is diagonal and \(\tilde{g}_{mn} = S_m^{-1} V_{mn} S_n\). Each block can now be represented in the basis of its energy eigenstates. To reduce the number of degrees of freedom we impose a cutoff and include only basis states within an energy range \(E_a < E < E_b\). This results in a lower rank Hamiltonian which we call the reduced Hamiltonian

\[
\mathcal{G}(\vec{k}) = \begin{bmatrix}
G_0 & g_{01} & 0 & \cdots & 0 & g_{0,n} \\
g_{01}^\dagger & G_1 & g_{12} & \cdots & 0 & 0 \\
0 & g_{12}^\dagger & G_2 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & G_{n-1} & g_{n-1,n} \\
g_{0,n}^\dagger & 0 & 0 & \cdots & g_{n-1,n}^\dagger & G_n
\end{bmatrix}
\]  

(5)

The size of the reduced Hamiltonian depends on both block sizes and the energy cutoff, and the interplay between these two parameters will also affect the accuracy.

3. Numerical results

To examine the trade-offs between efficiency and accuracy for our method we tested it on a \([011]\)-oriented of GaAs/AlAs quantum well. This material choice eliminates the complications from strain. Different quantum well thicknesses were considered, but the supercell size was kept fixed at 512 unit cells in the \(z\)-direction. Periodic boundary conditions were imposed in all directions and the in-plane wave vector was set to \(k_x = k_y = 0\). We used the \(sp^2d^x s^*\) model with parameters
Figure 5. Band gaps and their respective absolute errors for GaAs/AlAs quantum wells as a function of the number of blocks and the energy cutoff ranges. The quantum well width was varied between 10 and 20 unit cells, as indicated in the figures. With a cutoff range of $-0.5$ to 9 eV it is possible to use 32 blocks of size 16 unit cells without significant loss of accuracy for all the quantum well widths considered.

from reference [25], but neglecting spin–orbit coupling. This model uses the two-center approximation in the Slater–Koster formulation [11]. A compact expression of Slater–Koster formulation can be found in reference [26].

Hopping parameters are uniquely determined by the atoms being connected, but on-site energies can be ambiguous at interfaces. For example, an As atom at the interface may have both Ga and Al nearest neighbors. We linearly interpolate the on-site energies according to

$$E_{\text{m}}(\text{GaAs/AlAs}) = xE_{\text{m}}(\text{GaAs}) + (1-x)E_{\text{m}}(\text{AlAs}).$$

This should be accurate since the on-site energies in reference [25] are only weakly dependent on their neighboring atoms in different materials. We used a valence band offset of $E_v(\text{GaAs}) - E_v(\text{AlAs}) = 0.2$ eV. In figure 2 we show the band edge energies for a quantum well of GaAs in AlAs for different thicknesses. The band edge energies follow the expected pattern showing increasing quantum confinement with decreasing well thickness.

In figure 3 we present the calculated energies of our test system as a function of the number of blocks into which the supercell has been divided. With a well width of 12 unit cells and the number of unit cells per block being a power of two, the blocks are not commensurate with the well width and some blocks include both well and barrier material. The resulting energies, however, show little variation with the block size. With one block, the entire supercell is a block, and therefore the calculation is simply the unapproximated tight-binding model. This reference case gives the exact answer with which approximations will be compared. With 256 blocks each block contains two single zincblende unit cells, the least amount of coarse-graining shown in the figure. Since the supercell was fixed, an increasing number of blocks corresponds to a decreasing block size.

The energy cutoffs for the block basis states were varied for both the valence states (lower cutoff) and conduction states (upper cutoff). Three different energy ranges for the basis states of the blocks were considered: $-4$ eV to 16 eV, $-0.5$ eV to 13 eV, and $-0.5$ eV to 9 eV. Figure 3 shows that using up to 32 blocks (16 unit cells per block) both conduction and valence states are in good agreement with the exact tight-binding calculation for all three energy cutoffs.

The efficiency improvement from reducing the Hamiltonian was gauged by measuring the computation time, as shown in figure 4. The calculations initially become more efficient as the supercell is divided into more blocks, but the computation time eventually begins to increase again with a large number of small blocks. We find that by using eight blocks (containing 64 unit cells in each block) and an energy range of $-0.5$ eV to 13 eV, we get a substantial reduction in the computing time to about 5% of the full Hamiltonian with an error in the gap of less than 0.1%. We find the rank reduction of the Hamiltonian is larger than 60% and this value is relatively independent of the number of blocks. A slight increase in the Hamiltonian rank for large numbers of blocks accounts for the increasing time. We note that identical blocks are only diagonalized once and that some of the terms in the Hamiltonian representing interactions between blocks may also be identical and such terms are also computed only once. Reusing the basis states for blocks naturally decreases the computing time of our method. While our calculations were done on a single CPU workstation, our method lends itself to parallelization by diagonalizing different blocks on different cores.
Figure 6. The calculated DOS for a GaAs quantum well with AlAs barriers. The thickness of the quantum well was 20 unit cells and the barrier thickness was 246 unit cells. The energy cutoffs change the DOS to a small extent which can be seen in the figure. We find that using 32 blocks (16 unit cells per blocks) with energy cutoffs of $-0.5$ eV and 9.0 eV does not significantly affect the DOS near the band edges.

In figure 5 we show calculated gaps for quantum wells of different thicknesses as a function of the number of blocks and the energy cutoffs. The figure shows that the accuracy is insensitive to the quantum well thickness but the errors are smaller for larger cutoffs. The correlation between the energy cutoff and accuracy is not a simple linear relation, and for higher cutoffs the errors become less sensitive to the block size. If we reduce the high energy cutoff to 13 eV we can only use blocks up to size 32 unit cells before the accuracy begins to decrease. This behavior is expected since if the cutoffs were removed the original tight-binding model would be obtained and the errors would go to zero.

In figure 6 we show the calculated density of states (DOS) for a GaAs quantum well in AlAs. To calculate the DOS is more demanding than calculating the energy levels in a quantum well, since we need to sample many $k$-point in the Brillouin zone. In this case we used many cores for the calculation and we used an algorithm where only the eigenvalues were computed. For an energy cutoff of $(-0.5$ eV, 9.0 eV) and 16 unit cells per blocks we find a speedup of ten times compared with a calculation without truncation.

4. Summary

We have proposed a new method for constructing coarse-grained tight-binding models in a way that maintains the same degrees of freedom as the original tight-binding model. Systems with very different length scales could be modeled using different block sizes in different regions, and even leaving the original tight-binding model in those regions where atomic precision is required (e.g. near an impurity). We have tested this proposal by calculating energies for GaAs/AlAs quantum wells using an $sp^3d^5s^*$ tight-binding model. We coarse-grained the system using different block sizes to reduce the rank of the Hamiltonian, and found a significant reduction in computation time, as much as a factor of 10, with insignificant errors compared to the exact calculation. While we have tested the model using a 1D system, there is no special simplification that occurs in 1D. Our method is completely general and could be applied to systems in one, two and three dimensions. Our method inherits the existing limitations of tight-binding models. Parameters for a particular material do not transfer in an obvious way to a different material. For small systems with a surface nearby it is unclear if (or how) the parameters change. However our method does not introduce any new uncertainties since it only speeds up the calculations of an existing model.

Acknowledgments

This work was supported by NanoLund, the special overseas program of Wuhan University and the National Natural Science Foundation of China (Grant 91850207), the National Key R & D Program of China (Grant 2017FFA0303402 and Grant 2020YFA0211303).

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Tian-Xiang Liu © https://orcid.org/0000-0002-9210-0169
Mats-Erik Pistol © https://orcid.org/0000-0002-0234-3408
References

[1] Cohen M L and Bergstresser T K 1966 Band structures and pseudopotential form factors for fourteen semiconductors of the diamond and zinc-blende structures Phys. Rev. 141 789–96

[2] Chelikowsky J R and Cohen M L 1976 Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors Phys. Rev. B 14 556–82

[3] Yan Voon L L and Willatzen M 2009 The k · p Method: Electronic Properties of Semiconductors (Heidelberg: Springer)

[4] Jancu J-M, Scholz R, de Andrada e Silva E A and La Rocca G C 2005 Atomistic spin–orbit coupling and k · p parameters in III–V semiconductors Phys. Rev. B 72 193201

[5] Goringe C M, Bowler D R and Hernández E 1997 Tight-binding modelling of materials Rep. Prog. Phys. 60 1447–512

[6] Pryor C E and Pistol M-E 2015 Atomic k · p theory J. Appl. Phys. 118 225702

[7] Tan Y P, Povolotskyi M, Kubis T, Boykin T B and Klimeck G 2015 Tight-binding analysis of Si and GaAs ultrathin bodies with subatomic wave-function resolution Phys. Rev. B 92 085301

[8] Tan Y, Povolotskyi M, Kubis T, Boykin T B and Klimeck G 2016 Transferable tight-binding model for strained group IV and III–V materials and heterostructures Phys. Rev. B 94 045311

[9] Persson M P and Xu H Q 2006 Electronic structure of free-standing GaAs/AlGaAs nanowire superlattices Phys. Rev. B 73 035328

[10] Natan A, Benjamini A, Naveh D, Kronik L, Tiago M L, Beckman S R and Chelikowsky J R 2008 Real-space pseudopotential method for first principles calculations of general periodic and partially periodic systems Phys. Rev. B 78 075109

[11] Slater J C and Koster G F 1954 Simplified LCAO method for the periodic potential problem Phys. Rev. 94 1498–524

[12] Chadi D J and Cohen M L 1975 Tight-binding calculations of the valence bands of diamond and zincblende crystals Phys. Status Solidi b 68 405–19

[13] Papaconstantopoulos D A and Melh M J 2003 The Slater Koster tight-binding method: a computationally efficient and accurate approach J. Phys.: Condens. Matter. 15 R413–40

[14] Vogl P, Hjalmarson H P and Dow J D 1983 A semi-empirical tight-binding theory of the electronic structure of semiconductors J. Phys. Chem. Solids 44 365–78

[15] Yu P Y and Cardona M 2005 Fundamentals of Semiconductors: Physics and Materials Properties 3rd edn (Heidelberg: Springer)

[16] Sutton A P, Finnis M W, Pettifor D G and Ohta Y 1988 The tight-binding bond model J. Phys. C: Solid State Phys. 21 35–66

[17] Luisier M, Schenk A, Fichtner W and Klimeck G 2006 Atomistic simulation of nanowires in the sp³d⁵s* tight-binding formalism: from boundary conditions to strain calculations Phys. Rev. B 74 205323

[18] Boykin T B, Bowen R C and Klimeck G 2001 Electromagnetic coupling and gauge invariance in the empirical tight-binding method Phys. Rev. B 63 245314

[19] He Y, Tan Y, Jiang Z, Povolotskyi M, Klimeck G and Kubis T 2016 Surface passivation in empirical tight binding IEEE Trans. Electron Devices 63 954–8

[20] Chang Y C and Schulman J N 1982 Theory of heterostructures: a reduced Hamiltonian method with evanescent states and transfer matrices J. Vac. Sci. Technol. 21 540–3

[21] Schulman J N and Chang Y-C 1983 Reduced Hamiltonian method for solving the tight-binding model of interfaces Phys. Rev. B 27 2346–54

[22] Cardy J, Goddard P and Yeomans J 1996 Scaling and Renormalization in Statistical Physics (Cambridge Lecture Notes in Physics) (Cambridge: Cambridge University Press)

[23] Zinn-Justin J 2007 Phase Transitions and Renormalization Group (Oxford Graduate Texts) (Oxford: Oxford University Press)

[24] Löwdin P O 1950 On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals J. Chem. Phys. 18 365–75

[25] Jancu J-M, Scholz R, Beltram F and Bassani F 1998 Empirical spds* tight-binding calculation for cubic semiconductors: general method and material parameters Phys. Rev. B 57 6493–507

[26] Podolsky A V and Vogl P 2004 Compact expression for the angular dependence of tight-binding Hamiltonian matrix elements Phys. Rev. B 69 235101