Self-consistent tight-binding method

Feng Liu
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831
(Received 18 May 1995)

A self-consistent tight-binding formalism is described. The self-consistency is achieved by the introduction of a chemical hardness matrix and a generalization of the Hückel model to make the tight-binding Hamiltonian an implicit functional of the charge density. Studies of the band structures of diamond and face-centered cubic Si demonstrate that the method has very good transferability and shows promise for applications to systems involving large charge transfer. Also, the method can easily be implemented to study spin-polarized systems.

The tight-binding (TB) method is one of the most widely used methods in theoretical solid-state physics and materials science. It has been applied, for example, to study the band structures of semiconductors[1–4] and the magnetic properties of transition metals.[5] Besides its computational efficiency, the TB method is a very useful tool to study electronic structure because first-principles techniques based on the local-density approximation (LDA) usually underestimate the band gaps of semiconductors.[6,7] Another important application of the TB method is tight-binding total-energy (TBTE) calculations.[8–15] The recent development of the order-N density-matrix algorithm[16,17] makes the TBTE approach even more promising for the future. So far, most TB studies have been focused on monatomic systems It is quite difficult to extend TB studies to binary systems, especially to those polar systems involving large charge transfer, because the parameter fitting involved is much more complex and ambiguous. Drawbacks of the TB approach that have been noted in the literature include transferability and self-consistency. Both of these questions are addressed in this paper.

We propose a self-consistent tight-binding (STB) scheme to circumvent some of the difficulties related to non-self-consistency. In this approach, with the introduction of a chemical hardness matrix and a generalization of the Hückel model, the one-electron TB Hamiltonian is self-consistently determined according to the charge distribution so that the effect of charge transfer is explicitly taken into account. The method is expressed in terms of nonorthogonal bases. Most input parameters, such as atomic eigenvalues, chemical hardness matrices, and atomic basis functions can be obtained directly from *ab initio* atomic calculations. The scaling functions used in the generalized Hückel model can be fitted to the first-principles band structure; this is the only fitting involved. The method has been applied to calculate the band structures of silicon in the diamond and face-centered cubic (fcc) structures. The scaling function was first adjusted for the diamond structure. Not only the valence bands but also the two lowest conduction bands obtained are in excellent agreement with the first-principles results, after adjusting the latter to give the correct band gap. Using the same scaling function, the bands of fc silicon were then calculated. The results are again in rather good agreement with the first-principles calculations, showing the good transferability of the method.

We begin the discussion with the definition of the chemical hardness matrix. The total electronic energy of an atom can be expressed as a functional of the occupation numbers of the electronic states,

$$E_t = E\{n_i\} \ . \tag{1}$$

Assuming the occupation numbers n_i to be continuous variables, E_t can be expanded in a Taylor series,

$$E_t = E_0 + \sum_i \frac{\partial E}{\partial n_i}\Delta n_i + \frac{1}{2}\sum_{ij}\frac{\partial^2 E}{\partial n_i \partial n_j}\Delta n_i \Delta n_j + \frac{1}{6}\sum_{ijk}\frac{\partial^3 E}{\partial n_i \partial n_j \partial n_k}\Delta n_i \Delta n_j \Delta n_k + \cdots \ . \tag{2}$$

The eigenvalue of an electronic state is defined as

$$\epsilon_i = \frac{\partial E}{\partial n_i} \ . \tag{3}$$

Then we can show that

$$\epsilon_i = \epsilon_i^0 + \sum_j \frac{\partial \epsilon_i}{\partial n_j}\Delta n_j + \frac{1}{2}\sum_{jk}\frac{\partial^2 \epsilon_i}{\partial n_j \partial n_k}\Delta n_j \Delta n_k = \epsilon_i^0 + \sum_j \eta_{ij}^{(1)}\Delta n_j + \frac{1}{2}\sum_{jk}\eta_{ijk}^{(2)}\Delta n_j \Delta n_k \ , \tag{4}$$

where we have defined

$$\eta_{ij}^{(1)} = \frac{\partial \epsilon_i}{\partial n_j} \tag{5}$$

and

$$\eta_{ijk}^{(2)} = \frac{\partial^2 \epsilon_i}{\partial n_j \partial n_k} \tag{6}$$

as the first- and second-order chemical hardness matrices. Higher-order terms are neglected in Eq. (4).

The concept of the chemical hardness matrix was introduced by Parr and Pearson.[18] Physically, it describes how the chemical potential of an isolated system will change in response to a change in the number of electrons. More recently, Teter has used the chemical hardness matrix as an additional condition to generate norm-conserving pseudopotentials.[19] He showed that the

transferability of the pseudopotential could be improved substantially with this additional condition, which implies that the chemical hardness matrix is an intrinsic atomic property that represents the change of the eigenvalue in response to the change in the electronic configuration. In the following, we will show that the chemical hardness matrix can readily be incorporated into the TB formalism to determine the diagonal elements of the Hamiltonian. In combination with a generalized Hückel theory, the whole Hamiltonian can then be determined self-consistently.

The TB method is generally formulated within the linear-combination-of-atomic-orbitals (LCAO) approximation. The total charge density of a molecular or solid is decomposed into one-center contributions of effective atomic density. The renormalized atomic density will redefine the effective atomic eigenvalues in the new chemical environment. In the STB scheme, instead of treating the diagonal elements of the TB Hamiltonian as adjustable parameters in the conventional way, we set the diagonal elements equal to the effective atomic eigenvalues, namely,

$$H_{ii}=\epsilon_i \ , \tag{7}$$

which can be calculated directly from Eq. (4). As a result, this part of the Hamiltonian becomes an implicit functional of charge density since the effective atomic occupation number n_i in a molecule or solid represents the integrated charge density in the subspace of orbital i. For the off-diagonal elements, we adopt a generalized Hückel model to scale them in accordance with the diagonal terms as the following:

$$H_{ij}=K_{ij}(H_{ii}+H_{jj})S_{ij} \ , \tag{8}$$

where S_{ij} is the overlap matrix element and K_{ij} is a scaling function that will be discussed in detail later. In this way, the Hamiltonian has to be found self-consistently because it depends on the integrals of the charge densities of the orbitals, i.e., the occupation numbers n_i. In each step, n_i can be obtained either from Mulliken analysis[20] if the Hamiltonian is solved by direct diagonalization or from the diagonal terms of the density matrix if the Hamiltonian is solved by the density-matrix algorithm.[16,17] For large systems, the latter approach is highly recommended because of its superior linear-scaling property.

We use Si as an example to demonstrate the validity of the method by calculating the band structures of the diamond and fcc structures. First, to obtain the reference atomic eigenvalues [ϵ_i^0 in Eq. (4)], we carried out LDA atomic calculations at the ground-state electronic configuration with $n_s=2$ and $n_p=2$, where n_s and n_p denote the occupation numbers of the s and p levels, respectively. To obtain the atomic eigenvalues as a functional of occupation number, we did a series of atomic calculations with n_s and n_p varying from 0 to 4. The results were fitted to Eq. (4) to extract the first- and second-order chemical hardness matrices as the following, in units of eV/electron and eV/electron^2, respectively:

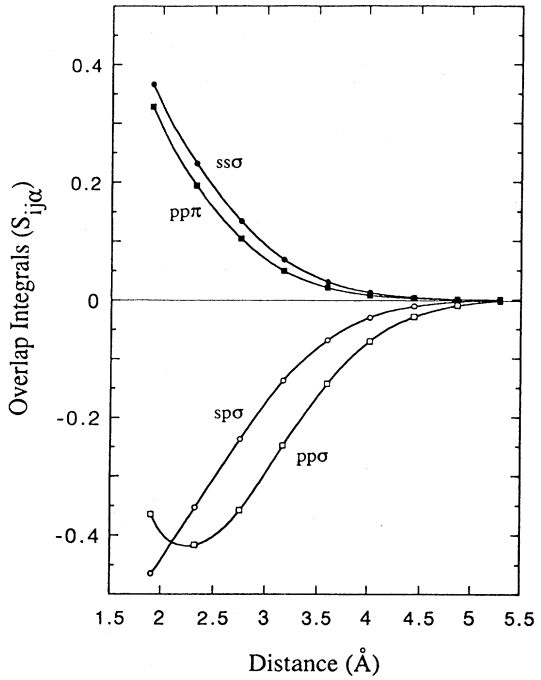

FIG. 1. Slater-Koster overlap integrals as a function of interatomic distance. Symbols are data from the dimer calculations and smooth lines are least-square fits to the data using Eq. (11).

$$\eta^{(1)}=\begin{bmatrix} 8.343 & 7.423 \\ 7.550 & 6.808 \end{bmatrix} , \tag{9}$$

$$\eta^{(2)}=\begin{bmatrix} -1.210 & -3.264 & -1.462 \\ -1.314 & -3.360 & -1.440 \end{bmatrix} . \tag{10}$$

Next, we did a number of Si-dimer calculations, using the discrete variational LCAO–molecular-orbital method.[5,21] The overlap matrix elements were decomposed into Slater-Koster overlap integrals ($S_{ss\sigma}$,$S_{sp\sigma}$,$S_{pp\sigma}$,$S_{pp\pi}$) within the two-center approximation.[22] In Fig. 1 we plot the resultant overlap integrals as a function of interatomic distance r. As one can see, all of the overlap integrals decrease monotonically in magnitude as the interatomic distance increases, except $S_{pp\sigma}$, which turns up at a separation of 2.3 Å. Also, all of

TABLE I. Coefficients used in Eq. (11) for calculating Slater-Koster overlap integrals.

$ij\alpha$	a_0	a_1	a_2	a_3	a_4	a_5
$ss\sigma$	1.9021	87.821	−132.90	76.515	−16.518	1.1826
$sp\sigma$	2.6300	1465.8	−2637.0	1756.3	−510.22	48.720
$pp\sigma$	2.4340	2082.4	−3631.7	2352.2	−656.86	60.977
$pp\pi$	2.1277	117.48	−178.20	101.91	−21.479	1.4831

the overlap integrals die out at about 5 Å, so we used a cutoff of 5 Å for the matrix elements in the solid-state STB calculations; this corresponds to the inclusion of up to the third-nearest neighbors in both the diamond and fcc structures. The smooth curves in Fig. 1 are least-square fits to the calculated data with the functional form

$$S_{ij\alpha}=e^{-a_0 r}(a_1+a_2 r+a_3 r^2+a_4 r^3+a_5 r^4)\ . \qquad (11)$$

These fits were used in the solid-state STB calculations. The fitting coefficients are tabulated in Table I.

From the dimer calculations, using Eq. (8), we also calculated the scaling factors K_{ij} in the Hückel model from the elements of the Hamiltonian and the overlap matrix. In general, these scaling factors were found to increase with increasing separation rather than remaining constant. We therefore chose to use a generalized Hückel model with distance-dependent scaling functions of the form

$$K_{ij}=e^{a_0 r}(a_1+a_2 r+a_3 r^2)\ .$$

We first calculated the band structure of diamond-cubic Si with all of the input parameters obtained from atomic and diatomic calculations with no further fitting. The results, however, were not satisfactory. For example, instead of an indirect band gap of 1.1 eV, we obtained a direct gap of 1.2 eV at the Γ point. We therefore repeated the calculations, adjusting the scaling functions K_{ij} until the best bands were achieved. The optimized coefficients are listed in Table II.

We started the STB calculations of the band structures of diamond and fcc Si by setting up the initial Hamiltonian with the ground-state atomic eigenvalues ($H_{ss}=\epsilon_s^0=-10.86$ eV and $H_{pp}=\epsilon_p^0=-4.2$ eV) and occupation numbers ($n_s=2$, $n_p=2$). In both cases, it only took a few cycles (~ 4) for the Hamiltonian to converge. The converged effective atomic electronic configurations are $n_s=1.416$ and $n_p=2.584$ for the diamond structure and $n_s=1.767$ and $n_p=2.233$ for the fcc structure. In Fig. 2, the STB band structure of diamond-cubic Si at the experimental lattice constant of 5.4 Å is plotted in comparison to the LDA band structure obtained from the pseudopotential plane-wave calculation. The top of the valence bands at the Γ point has been aligned for the two calculations. The LDA conduction bands have been shifted upward rigidly by 0.60 eV relative to the valence bands to match the experimental band gap. Also, we have intentionally made the STB valence-band width slightly larger than the LDA bandwidth to agree with the experimental value. The dispersion of all of the STB

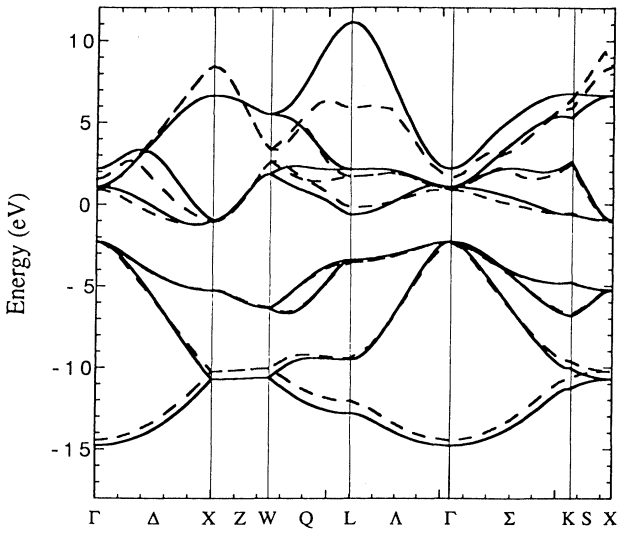

FIG. 2. Band structure of diamond-cubic Si at the equilibrium lattice constant. Solid lines are STB results and dashed lines are LDA results. The top of the valence bands has been aligned for the two calculations. The LDA conduction bands have been rigidly shifted up by 0.60 eV.

valence bands as well as the two lowest conduction bands agrees very well with the LDA results. The rms deviation of these bands is 0.11 eV, which is about the same as the best that has previously been achieved.[1] The agreement for the two topmost conduction bands is not as good, possibly because these two high-energy bands involve contributions from $3d$ and $4s$ states.[4] Using the same set of input parameters, we then applied the STB method to calculate the band structure of fcc Si at the lattice constant of 3.8 Å, which is theoretically predicted from LDA pseudopotential total-energy calculations. The results are shown in Fig. 3. Again, the agreement between the STB and LDA bands in the energy region up to a few eV above the Fermi energy is fairly good, with a rms error of 0.46 eV up to the Fermi energy. Allen *et al.*[4] showed that reasonably transferable nonorthogonal Slater-Koster parameters could be obtained within a group of hypothetical structures of Si at the constant density of the diamond structure by fitting them simultaneously to the first-principles band structures of the various members. Here we are able to show that the STB parameters adjusted for the diamond structure can produce fairly good results for the fcc structure. The fact that the two systems studied here have very different

TABLE II. Coefficients used in Eq. (12) for calculating scaling functions of the generalized Hückel model.

$ij\alpha$	a_0	a_1	a_2	a_3
$ss\sigma$	0.468 84	0.641 76	−0.230 55	0.026 428
$sp\sigma$	0.595 60	0.481 50	−0.173 16	0.020 419
$pp\sigma$	0.578 09	0.499 66	−0.185 36	0.024 962
$pp\pi$	0.619 51	0.630 20	−0.232 98	0.027 049

coordination numbers (4 vs 12) and quite different atomic densities (0.051 vs 0.073 atoms/Å^3) demonstrates the good transferability of the STB method.

We would like to point out that the chemical hardness matrix introduced here can be easily implemented in any existing tight-binding scheme as a way to readjust the on-site diagonal elements to improve the transferability. It can be applied not only to the band-structure calculations but also to the TBTE calculations. Tománek and Schlüter[9] and Goodwin *et al.*[10] have shown that to use the TBTE scheme designed for the bulk phase to study clusters one has to add an additional term to the total energy to treat the intra-atomic Coulomb interactions. The Coulomb repulsion parameter used in the term can be considered, in a way, as a simplified first-order chemical hardness matrix with all diagonal elements equal to the Coulomb repulsion energy and all off-diagonal elements equal to zero. It treats the effect of net charge transfer between atoms in an approximate way but ignores the effect of charge promotion among the electronic levels within the atom. More recently, Mercer and Chou[14] introduced an intra-atomic matrix to change the on-site energies according to the local coordination and geometry in order to eliminate the structure-dependent function in the total-energy expression. Cohen *et al.*[15] devised a TB approach with on-site terms varying as a function of local "density" to improve the transferability for metallic systems.

In conclusion, we have described the formalism of a self-consistent tight-binding method. Two key ingredients of this method are the introduction of the chemical hardness matrix and the generalization of the Hückel model. The method has several advantages over the conventional non-self-consistent approach. First, as demonstrated by the studies of the band structures of diamond-cubic and fcc Si, the self-consistency leads to good transferability. Second, the method should be especially useful for systems involving large charge transfer, because the Hamiltonian is a functional of the charge distribution and the effect of charge transfer is included in a self-consistent manner. Third, many input parameters can be directly obtained from atomic, diatomic, and molecular calculations, so the amount of fitting work is reduced. Finally, the method can easily be extended to study magnetic problems by employing a spin-dependent chemical hardness matrix which can be obtained from spin-polarized atomic calculations.

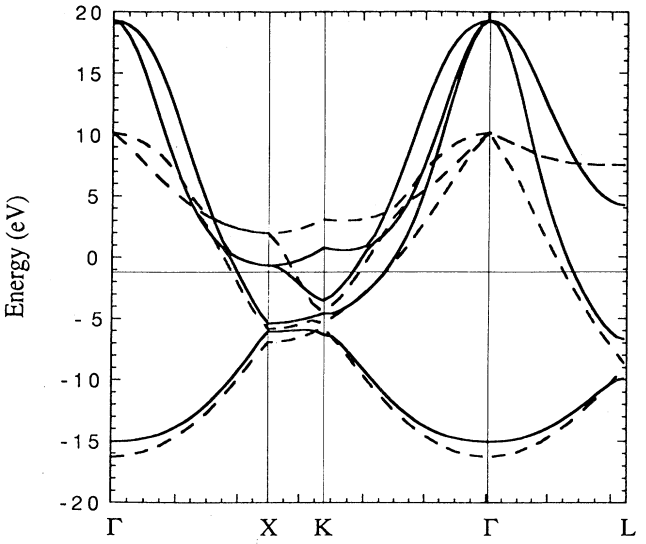

FIG. 3. Band structures of fcc Si at the equilibrium lattice constant. Solid lines are STB results and dashed lines are LDA results. The horizontal line at 1.1 eV marks the position of the Fermi energy, which is aligned for the STB and LDA bands.

The author is grateful for helpful discussions with David Vanderbilt, Mark Mostoller, Theodore Kaplan, and J. F. Cooke. Oak Ridge National Laboratory is operated for the U.S. Department of Energy by Martin Marieta Energy Systems, Inc. under Contract No. DE-AC05-84OR21400.

[1]L. F. Mattheiss and J. R. Patel, Phys. Rev. B **23**, 5384 (1981), and references therein.
[2]P. Vogl, H. J. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids **44**, 365 (1983).
[3]D. A. Papaconstantopoulos, *Handbook of the Band Structure of Elemental Solids* (Plenum, New York, 1986).
[4]Philip B. Allen, Jeremy Q. Broughton, and A. K. McMahan, Phys. Rev. B **34**, 859 (1986).
[5]Feng Liu, M. R. Press, S. N. Khanna, and P. Jena, Phys. Rev. B **38**, 5760 (1988); **39**, 6914 (1989).
[6]J. R. K. Bigger, D. A. McInnes, A. P. Sutton, M. C. Payne, I. Stich, R. D. King-Smith, D. M. Bird, and L. J. Clarke, Phys. Rev. Lett. **69**, 2224 (1992).
[7]Feng Liu, M. Mostoller, V. Milman, M. F. Chishlom, and T. Kaplan, Phys. Rev. B **51**, 17 192 (1995).
[8]D. J. Chadi, Phys. Rev. B **29**, 785 (1984).
[9]D. Tománek and M. A. Schlüter, Phys. Rev. Lett. **56**, 1055 (1986).
[10]L. Goodwin, A. J. Skinner, and D. G. Pettifor, Europhys. Lett. **9**, 701 (1989).
[11]C. Z. Wang, C. T. Chan, and K. M. Ho, Phys. Rev. B **39**, 8586 (1989).
[12]A. Caro, S. R. de Debiaggi, and M. Victoria, Phys. Rev. B **41**, 913 (1990).
[13]M. Menon and K. R. Subbaswamy, Phys. Rev. B **47**, 12 754 (1993).
[14]J. L. Mercer, Jr. and M. Y. Chou, Phys. Rev. B **49**, 8506 (1994).
[15]R. E. Cohen, M. J. Mehl, and D. A. Papaconstantopoulos, Phys. Rev. B **50**, 14 694 (1994).
[16]X.-P. Li, R. W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10 891 (1993).
[17]M. S. Daw, Phys. Rev. B **47**, 10 895 (1993).
[18]R. G. Parr and R. G. Pearson, J. Am. Chem. Soc. **105**, 7512 (1983).
[19]M. Teter, Phys. Rev. B **48**, 5031 (1993).
[20]R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955); **23**, 1841 (1955).
[21]D. E. Ellis and G. S. Painter, Phys. Rev. B **2**, 2887 (1970).
[22]J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).