

Transferable nonorthogonal tight-binding scheme for silicon

Madhu Menon and K. R. Subbaswamy

Department of Physics and Astronomy, University of Kentucky, Lexington, Kentucky 40506

(Received 19 May 1994)

A minimal-parameter tight-binding theory incorporating explicit use of nonorthogonality of the basis is used to generate a transferable scheme for silicon. Good results are obtained for band structure, phase diagram, and bulk phonons. The diamond structure is found to be the ground state even when compared with the clathrate structure. The results for clusters show good agreement with *ab initio* predictions. The theory differs from the conventional orthogonal schemes in three main respects: (1) only three adjustable parameters are employed, (2) no artificial cutoff is used for interactions and, (3) the need for a coordination-dependent energy term has been obviated for clusters of any size.

I. INTRODUCTION

Recently, there has been a resurgence of interest in tight-binding schemes for semiconductors in general, and Si in particular.¹⁻⁵ This comes, in part, from the interest in heterostructures, amorphous Si, the clathrate form, and other forms where *ab initio* and local-density approximation (LDA) schemes prove inadequate due to system size considerations. The goal is to obtain a transferable scheme that works well in the range all the way from a few atoms to the condensed solid. The published schemes range from that of Chadi and Cohen⁶ with eight parameters to that of Mercer and Chou⁴ with more than 40 parameters. The former is designed to get a good description of the band structure in the diamond phase, but shows limited transferability when applied to clusters.³ The latter, while giving a good phase diagram by including intra-atomic parameters with distance dependence, has not been tested for clusters of arbitrary sizes. Also, none of the orthogonal schemes have reported comparison between the clathrate structure⁷ and diamond in their phase diagrams; this is a stringent test, since they are both fourfold coordinated networks.

Harrison⁸ attempted to provide a minimal tight-binding theory with just four parameters (in addition to four dimensionless universal constants) that could describe qualitatively a wide range of materials and properties. While the focus of his work was tetrahedral solids, in later work with van Schilfgaarde^{9,10} he emphasized the necessity of including the nonorthogonality of the local environment in multicoordinated structures. This important factor has been generally overlooked by those seeking a transferable scheme. Nonorthogonal tight-binding parameters for accurately reproducing band structures for various crystalline phases of silicon have been obtained by Mattheiss and Patel¹¹ and by Allen *et al.*¹² These works not only involve multiple parameters to fit the Hamiltonian and overlap matrix elements but also assume the volume to be constant and, therefore, cannot be used in molecular dynamics simulations and for clusters.

We have demonstrated earlier in our tight-binding molecular dynamics work on silicon³ and carbon¹³ clusters that including nonorthogonality in Harrison's tight-binding scheme markedly improves the agreement with *ab initio* results¹⁴⁻¹⁶ for clusters in the range up to $N = 10$. Our scheme was further improved by Ordejón *et al.*¹⁷ We note here that clusters in the range 6-10 are particularly challenging because of high coordination.

In this paper, we show that the nonorthogonal tight-binding theory gives a transferable scheme with only three *adjustable* parameters, giving good agreement with experiment or *ab initio* work for (i) the zero-temperature phase diagram (including the clathrate structure), (ii) valence band structure and phonon dispersion for the crystal, and (iii) structure and cohesive energies and frequencies of small clusters.

None of the other tight-binding schemes have so few adjustable parameters and none have demonstrated applicability to these many properties. In the following section, we describe our tight-binding scheme and the computed results. In the final section, the scheme is compared and contrasted with other methods.

II. TECHNIQUE

The details of the technique as applied to clusters of atoms are given in Refs 3 and 17. Here we give a brief summary and show the application of the technique to solids. The total energy is given by the sum

$$U = U_{\text{el}} + U_{\text{rep}} + U_0, \quad (1)$$

where U_{el} is the sum of the one-electron energies ϵ_k for the occupied states:

$$U_{\text{el}} = \sum_k^{\text{occ}} \epsilon_k, \quad (2)$$

and U_{rep} is given by a repulsive pair potential

$$U_{\text{rep}} = \sum_i \sum_{j>i} \chi(r_{ij}). \quad (3)$$

U_0 is constant that merely shifts the zero of energy. The nonorthogonal tight-binding scheme differs from the conventional schemes in the evaluation of the one-electron energies in Eq. 2. In conventional schemes an unknown orthonormal set of basis is used in the construction of the matrix elements of the Hamiltonian. In the generalized scheme the wave function of the system is given in terms of the nonorthogonal basis as

$$\psi_n = \sum_a c_a^n \phi_a. \quad (4)$$

The characteristic equation then becomes

$$\sum_j (H_{ij} - E_n S_{ij}) c_j^n = 0, \quad (5)$$

where

$$H_{ij} = \int \phi_i^* H \phi_j d^3r, \quad (6)$$

and

$$S_{ij} = \int \phi_i^* \phi_j d^3r. \quad (7)$$

In matrix form Eq. (5) becomes

$$(\mathbf{H} - E_n \mathbf{S}) \mathbf{C}^n = 0. \quad (8)$$

If V_{ij} denote the Hamiltonian matrix elements in orthogonal theory, the Hamiltonian matrix elements in the nonorthogonal scheme is constructed according to the prescription given by van Schilfgaarde and Harrison,⁹

$$H_{ij} = V_{ij} \left[1 + \frac{1}{K} - S_2^2 \right], \quad (9)$$

where

$$S_2 = \frac{(S_{ss\sigma} - 2\sqrt{3}S_{sp\sigma} - 3S_{pp\sigma} + 3S_{pp\pi})}{4} \quad (10)$$

is the nonorthogonality between sp^3 hybrids and K is a nonorthogonality coefficient. Note that unlike van Schilfgaarde and Harrison,⁹ we retain $S_{pp\pi}$ in S_2 for consistency since we set $S_{pp\pi}$ proportional to $V_{pp\pi}$. We find that by doing so the pairwise repulsive coefficient can be significantly reduced and perhaps eliminated. The quantities $S_{\lambda\lambda'\mu}$ in turn are determined from

$$S_{\lambda\lambda'\mu} = \frac{2V_{\lambda\lambda'\mu}}{K(\epsilon_\lambda + \epsilon_{\lambda'})}. \quad (11)$$

In this form H_{ij} and S_{ij} are more transferable between different environments than the original orthogonal matrix elements V_{ij} , since the effect of overlaps in different environments is explicitly contained. In our earlier work,³ S_2 was held to be distance independent. The $V_{\lambda\lambda'\mu}(d)$ are taken to have a simple exponential dependence with d :

$$V_{\lambda\lambda'\mu}(d) = V_{\lambda\lambda'\mu}(d_0)e^{-\alpha(d-d_0)}, \quad (12)$$

where d_0 is the bond length for the crystal and α is a pa-

rameter whose value is to be determined. Later Ordejón *et al.*,¹⁷ showed that putting the same distance scaling as Eq. (12) into S_2 gives much better results for energies and obviates the need for a coordination-dependent term, introduced by Tomañek and Schluter¹⁸ and used in the conventional tight-binding molecular dynamics schemes.¹⁹ The elimination of the coordination-dependent term is crucial for transferability.

The repulsive term is taken to be a sum of classical short-ranged pair potentials, also with a simple exponential dependence with distance

$$\chi(r) = \chi_0 e^{-\beta(r-d_0)}. \quad (13)$$

We fix $\beta=4\alpha$. In our nonorthogonal theory this pairwise repulsive term is very small, and can in fact be omitted altogether.

We further introduce a simple distance dependence in the nonorthogonality coefficient, K ,

$$K(r) = K_0 + \alpha(r - d_0)^2. \quad (14)$$

As shown later, this dependence yields better transferability by giving improved agreement at both the cluster and bulk ends.

The eigenvalues of a system with nonorthogonal basis set can then be obtained from

$$\det |H_{ij} - E S_{ij}| = 0. \quad (15)$$

Evaluation of (15) is expedited by the use of the well known Cholesky factorization in which \mathbf{S} is factored into

$$\mathbf{S} = \mathbf{B}\mathbf{B}^\dagger. \quad (16)$$

This factorization is always possible provided \mathbf{S} is positive definite.

The theory involves the following six parameters: The diagonal matrix elements (atomic term values) ϵ_s and ϵ_p ; the covalent radius ($= d_0/2$); the interaction falloff rate α ; the nonorthogonality constant K ; and the repulsive coefficient χ_0 . Of these, ϵ_s , ϵ_p , and d_0 are set *a priori* from Harrison's work⁸, and are not adjusted. The parameter values are summarized in Table I. The four $V_{\lambda\lambda'\mu}$ are derived from the dimensionless *universal* parameters through a prescription given by Harrison,⁸ and have the values $V_{ss\sigma} = -2.37$ eV, $V_{sp\sigma} = 2.52$ eV, $V_{pp\sigma} = 3.32$ eV, $V_{pp\pi} = -1.07$ eV.

Thus, in our scheme there are only three *adjustable* parameters, α , K , and χ_0 . These can be fitted simply either to a dimer or the crystalline solid for experimental bond length and frequencies. As will be shown, fitting at either end does not substantially alter results at the other end owing to the transferability. In our case, we fixed the parameters to minimize the errors at both ends.

TABLE I. Parameters used in the present scheme for silicon.

<i>A priori</i> parameters			Adjustable parameters		
ϵ_s (eV)	ϵ_p (eV)	d_0 (Å)	α (Å ⁻¹)	K_0	χ_0 (eV)
-13.55	-6.52	2.36	1.6	1.7	0.05

III. RESULTS

A. Solid

In this section, we present our results obtained using the nonorthogonal scheme. All the results are obtained without introducing any artificial cutoff in the interactions. In practice this is achieved by increasing the shell size until results remain unchanged. We find inclusion of up to the third neighbor shell (corresponding to a cutoff of 5 Å) to be sufficient (and in some cases, e.g., band structure calculation, necessary) for obtaining converged results.

Figure 1 shows the zero-temperature phase diagram for silicon obtained by our method. In computing the electronic energies special point integrations were performed and tested for convergence. Compared to the higher coordinated structure (fcc, sc, β tin) the diamond structure is lower in energy, as is found in other transferable schemes as well.^{1,4,5} As can be seen in the figure, in our scheme the diamond structure is lowest in energy even when compared with the clathrate structure⁷ with the same coordination, although the difference is quite small (0.04 eV/atom). We believe inclusion of nonorthogonality is crucial in correctly differentiating these two structures which are so close energetically.

In Fig. 2 we show the band structure for the diamond structure. This is almost identical to the band structure obtained by Harrison⁸ using the universal param-

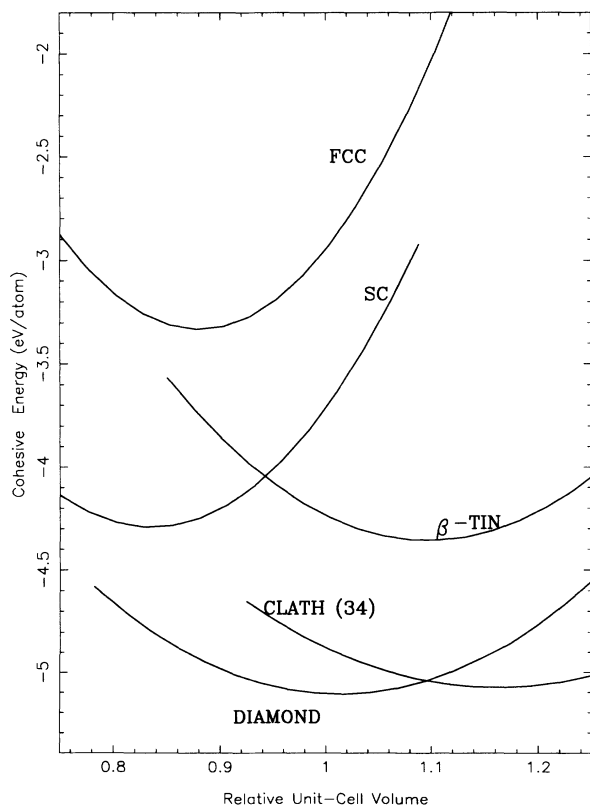


FIG. 1. The cohesive energies of various structures of silicon using the present scheme.

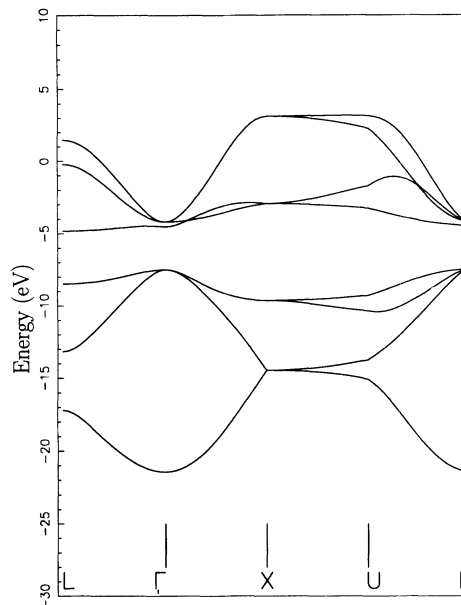


FIG. 2. The band structure for crystalline silicon in the diamond structure.

eters within the orthogonal scheme, although we obtain an indirect gap. As can be seen, the valence band is well reproduced, while the gap is larger than experiment. This is a consequence of the minimal basis set, and is true for all such tight-binding schemes. One can easily improve the conduction bands by the inclusion of s^* states as used in later work by Harrison.²⁰ These states, however, are not important for determining ground state energies and forces.

The force constants were evaluated by computing energy changes due to mutual displacements. The vibrational frequencies at some symmetry points are given in Table II. While the acoustic branch phonons are in excellent agreement with experiment, the optical modes show deviations up to about 27% from experiment.²¹ This should be contrasted with the best orthogonal schemes that achieve agreement to only within 45% while using many more parameters.⁵ As shown in the next section on clusters, the same scheme gives a reasonable vibrational frequency for the symmetric stretch mode for the dimer. In Sec. IV we will indicate the possibility of further improvements at both the cluster and bulk ends.

B. Clusters

We next briefly describe our results for small silicon clusters using the present scheme. All geometries were optimized by molecular dynamics relaxation using precisely the same tight-binding parameters. The molecular dynamics procedure has been previously described.³ Once again, no cutoff is used in all our calculations. Although our method was originally introduced with applications to these clusters, the present model (which includes modifications proposed by Ordejón *et al.*¹⁷) gives overall improvement in the results. The most notable improvement is the absence of an *ad hoc* coordination-

TABLE II. Comparisons showing transferability for silicon. The numbers in parentheses are experimental values.

	Bond length (Å)	Binding energy (eV/atom)	Vibrational frequency (cm ⁻¹)
Si (dimer)	2.29 (2.24) ^a	1.57 (1.56) ^b	587 (517) ^c
Si (diamond)	2.35 (2.35) ^a	5.11 (4.92) ^d	$\Gamma_{TO}=656$ (517) ^c $L_{LO}=496$ (417) ^c $L_{LA}=408$ (368) ^c $X_{LO}=580$ (463) ^c $X_{TO}=448$ (414) ^c

^aReference 21.

^bReference 14.

^cReference 22.

^dReference 4.

dependent term needed in earlier schemes.^{3,18,19} We use the same energy expression [Eq. (1)] for comparing energies in our search for minimum energy geometries. In Table III we give binding energies of the most stable clusters. In comparing our *absolute* cohesive energies with *ab initio* results, a constant shift of 1 eV to our computed values brought all values into excellent agreement. While the structural results are the same as in Ref. 17, the energies are in slightly better agreement with *ab initio* values.

For Si₂, we obtain a dimer bond length of 2.29 Å and a vibrational frequency of 587 cm⁻¹. The corresponding experimental values are 2.24 Å and 517 cm⁻¹, respectively. In Table II we compare our results with experiment for bond lengths, cohesive energies, and vibrational frequencies at the dimer and bulk ends to illustrate the transferability of the present scheme.

The minimum energy structure for Si₃ is found to be an open triangle with C_{2v} symmetry. For N = 4, we find the stable structure to be a rhombus (D_{2h} symmetry). These results are in good agreement with *ab initio* calculations.¹⁴

For Si₅, the lowest energy configuration is found to be a strongly compressed trigonal bipyramid (D_{3h} symmetry) with apex atoms holding the triangle together, shown in Fig. 3, in agreement with the *ab initio* results.¹⁴

TABLE III. Cohesive energies (in eV/atom) for Si_N clusters. Our computed values were shifted up uniformly by 1 eV to bring absolute values into agreement with *ab initio* values.

N	Symmetry	Binding energy (eV/atom)	
		<i>Ab initio</i> ^a	Present work
2		1.56	1.57
3	C _{2v}	2.54	2.49
4	D _{2h}	3.17	3.19
5	D _{3h}	3.3	3.36
6	C _{2v}	3.6	3.62
7	D _{5h}	3.8	3.78
8	C _{2h}	3.65	3.75
9	C _{3v}	3.6	3.80
10	C _{3v}	3.82	3.90
Solid		4.92 ^b	5.11

^aReference 15.

^bReference 4.

In the case of Si₆, we find the face-capped trigonal bipyramid (Fig. 3) to be 0.05 eV lower in energy over a distorted edge-capped trigonal bipyramid. This is in excellent agreement with *ab initio*¹⁴ results which find the former to be 0.04 eV lower in energy compared with the latter.

For Si₇, we find the pentagonal bipyramid (Fig. 3) to be more stable than a tricapped tetrahedron by 1.02 eV. This is in agreement with *ab initio* calculations,¹⁵ which give an energy difference of 0.952 eV.

In case of Si₈, the lowest energy structure is a distorted bicapped octahedron with C_{2h} symmetry (Fig. 3). Another minimum is obtained at 0.568 eV higher for an undistorted bicapped octahedron (D_{3h} symmetry). This is in excellent agreement with the *ab initio* calculations¹⁵ which find an energy difference of 0.56 eV between these two structures.

For Si₉, we have considered three geometries: the dis-

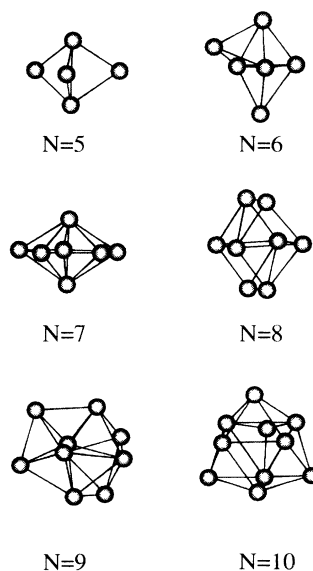


FIG. 3. Geometries of the lowest energy structures of small silicon clusters from N = 5 to 10 obtained using the present molecular dynamics scheme.

torted tricapped octahedron, a distorted tricapped trigonal prism proposed by Ordejón *et al.*¹⁷ (C_{2v} symmetry), and the tricapped octahedron with C_{3v} symmetry. We find the first structures to be unstable, distorting to the second structure. The tricapped octahedron with C_{2v} symmetry is found to be 1.08 eV less stable. The *ab initio* work of Raghavachari and co-workers¹⁵ only considered an undistorted trigonal prism when comparing with other structures and found all these structures to be almost isoenergetic when correlation effects were included. More recent *ab initio* calculations¹⁷ using an unrestricted Hartree-Fock calculation using 6-31G* basis including correlation effects have suggested the distorted tricapped trigonal prism to be slightly lower in energy than other structures. Approximate LDA calculations¹ also predict this structure to be lower in energy over the other two structures.¹⁷ In light of these findings this geometry must be considered as a candidate for the Si_9 cluster when searching for ground state by other methods.

For Si_{10} the minimum energy structure (a tetrapped trigonal prism with C_{3v} symmetry, Fig. 3) is 1.51 eV more stable than the tetrapped octahedron (T_d symmetry). This should be compared with the *ab initio* results which obtained an energy difference of 1.04 eV between these structures.

Transferability from solid to cluster is crucial to the study of crystal growth. Clusters in the range $N = 6-10$ are particularly challenging on account of the multicoordinated geometries predicted by *ab initio* calculations¹⁵ for these clusters. The orthogonal tight-binding schemes have found it necessary to introduce an *ad hoc* coordination-dependent term to get reasonable agreement with *ab initio* values for cohesive energies.^{18,19} This term is cutoff dependent and its use in molecular dynamics is rather awkward since it can give rise to spurious forces. This term also makes any accurate search for true ground states using simulated annealing methods impossible. We believe its presence is a necessary consequence of the lack of proper environment dependence present in conventional orthogonal tight-binding schemes.

IV. DISCUSSION AND SUMMARY

We have presented a simplified transferable generalized tight-binding scheme for silicon which incorporates the overlap interactions explicitly and obtained agreement on a wide range of properties for different phases. While previous attempts at constructing transferable tight-binding

schemes^{2,4,5} have had some success, the number of adjustable parameters had to be increased considerably. We believe this to be the result of the implicit assumption of orthogonality of the atomic basis in the formalism. Results for clusters in the range $N = 6-10$ presented here as in Ref. 17 clearly show that the coordination-dependent term found necessary in the orthogonal schemes is due to the lack of sufficient local coordination information in the electronic energy term. The present scheme employs no coordination-dependent term and by the proper treatment of the overlap interactions contains the effects of local atomic configuration entirely in the electronic part of Eq. (2). As shown by van Schilfhaarde and Harrison,^{9,10} the nonorthogonality can be separated out entirely as a two-body repulsion term only in systems dominated by two center bonds. The orthogonalization of atomic orbitals to approximate the Wannier functions, however, depends on the structure, and for multicoordinated systems the separation of nonorthogonality into a two-body repulsion term cannot be assumed.

In the present scheme, even with so few adjustable parameters, we have not exhausted all degrees of freedom. One can obtain a more accurate fit to LDA values for the various phases by employing a more sophisticated distance dependence in the parameters without altering cluster results. It is significant to note that Ordejón *et al.*¹⁷ used a very different distance dependence ($\alpha = 0.89 \text{ \AA}^{-1}$) while obtaining the same ground states for clusters in the range $N = 6-10$. Also, as mentioned before, the retention of $S_{pp\pi}$ in the expression for S_2 reduces the value of χ_0 needed (e.g., it is 0.486 eV in Ref. 17 while it reduces to 0.05 eV in the present work). Further, the results also did not change significantly when this small repulsive coefficient was eliminated altogether. Our goal in this work is merely to show that reasonable agreement with a wide range of properties is possible with the fewest parameters and a simple scaling of the parameters by explicitly treating the nonorthogonality in tight-binding theory.

ACKNOWLEDGMENTS

We are grateful to Richard Martin, Pablo Ordejón, Dmitry Lebedenko, Otto Sankey, Alexander Demkov, and Ernst Richter for helpful discussions. This research was supported in part by USDOE Contract No. DE-FC22-93PC93053, by NSF Grant No. EHR 91-08764, and by the University of Kentucky Center for Computational Sciences.

¹ O. F. Sankey and D. J. Niklewski, Phys. Rev. B **40**, 3979 (1989).

² L. Goodwin, A. J. Skinner, and D. G. Pettifor, Europhys. Lett. **9**, 701 (1989).

³ M. Menon and K. R. Subbaswamy, Phys. Rev. B **47**, 12 754 (1993).

⁴ J. L. Mercer, Jr. and M. Y. Chou, Phys. Rev. B **49**, 8506

(1994).

⁵ I. Kwon *et al.*, Phys. Rev. B **49**, 7242 (1994).

⁶ D. J. Chadi and M. L. Cohen, Phys. Status Solidi B **68**, 405 (1975).

⁷ G. B. Adams *et al.*, Phys. Rev. B **49**, 8048 (1994).

⁸ W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980), and references

- therein.
- ⁹ M. van Schilfgaarde and W. A. Harrison, *J. Phys. Chem. Solids* **46**, 1093 (1985).
- ¹⁰ M. van Schilfgaarde and W. A. Harrison, *Phys. Rev. B* **33**, 2653 (1986).
- ¹¹ L. F. Mattheiss and J. R. Patel, *Phys. Rev. B* **23**, 5384 (1981).
- ¹² P. B. Allen, J. Q. Broughton, and A. K. McMahan, *Phys. Rev. B* **34**, 859 (1986).
- ¹³ M. Menon, K.R. Subbaswamy, and M. Sawtarie, *Phys. Rev. B* **48**, 8398 (1993).
- ¹⁴ K. Raghavachari, *J. Chem. Phys.* **84**, 5672 (1986)
- ¹⁵ K. Raghavachari and C. M. Rohlfing, *J. Chem. Phys.* **89**, 2219 (1988).
- ¹⁶ K. Raghavachari and J. S. Binkley, *J. Chem. Phys.* **87**, 2191 (1987).
- ¹⁷ P. Ordejón, D. Lebedenko, and M. Menon, *Phys. Rev. B* **50**, 5645 (1994).
- ¹⁸ D. Tomañek and M. A. Schluter, *Phys. Rev. B* **36**, 1208 (1987).
- ¹⁹ K. Laasonen and R. M. Nieminen, *J. Phys. C* **2**, 1509 (1990).
- ²⁰ W. A. Harrison, *Phys. Rev. B* **24**, 5835 (1981).
- ²¹ G. Nilsson and G. Nelin, *Phys. Rev. B* **6**, 3777 (1972).
- ²² K. P. Huber and G. Herzberg, *Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).

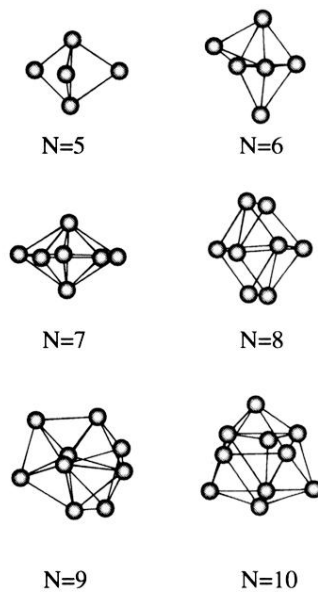


FIG. 3. Geometries of the lowest energy structures of small silicon clusters from $N = 5$ to 10 obtained using the present molecular dynamics scheme.