

# Quantum Monte Carlo Study of Silicon Self-interstitial Defects

W.-K. LEUNG<sup>a</sup>, R. J. NEEDS<sup>a,\*</sup>, G. RAJAGOPAL<sup>a</sup>, S. ITOH<sup>b</sup> and S. IHARA<sup>b</sup>

<sup>a</sup>*TCM Group, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK;*

<sup>b</sup>*Central Research Laboratory, Hitachi Ltd. (Japan), Kokubunji, Tokyo 185-8601*

We give a brief description of the variational and diffusion quantum Monte Carlo methods and their application to the study of self-interstitial defects in silicon. The diffusion quantum Monte Carlo calculations give formation energies for the most stable defects of about 4.9 eV, which is considerably larger than the values obtained in density functional theory methods. The quantum Monte Carlo results indicate a value for the formation+ migration energy of the self-interstitial contribution to self-diffusion of about 5 eV, which is consistent with the experimental data.

*Keywords:* Self-interstitials; Silicon; Diffusion quantum Monte Carlo; Self-diffusion; Density functional theory

## 1. INTRODUCTION

A sophisticated description of electron correlation is required to describe the structures and energetics of solids. Quantum Monte Carlo methods in the variational [1] (VMC) and diffusion [2, 3] (DMC) forms allow a direct assault on the many-body Schrödinger equation using statistical techniques which are capable of yielding highly accurate results. The great promise of these methods lies in the fact that electron-electron correlations are included explicitly, essentially without approximation, and the computational cost increases as the third power of the number of electrons, which is much more favourable than other correlated wave

function techniques and allows applications to large systems.

In this paper we give a brief introduction to the VMC and DMC methods and describe their application to the study of self-interstitial defects in silicon [4]. One of the important problems in the manufacture of sub-micron devices is the diffusion of dopant impurity atoms during thermal processing, which limits how small they can be made. To understand these effects requires a knowledge of diffusion on the microscopic scale in situations far from equilibrium. The diffusion of impurity atoms in silicon is controlled by intrinsic defects such as self-interstitials and vacancies, and it is therefore important to

---

\*Corresponding author. e-mail: rn11@phy.cam.ac.uk

improve our understanding of the behaviour of these defects.

Unfortunately it has not been possible to detect the self-interstitial directly, although their presence has been inferred from various measurements [5]. Measurements of the self-diffusion constant or self-diffusivity of silicon at high temperatures using radioactive isotopes of silicon as tracers have established an Arrhenius behaviour with an activation energy in the range 4.1–5.1 eV [6]. The self-diffusivity,  $D_{SD}$ , is usually written as the sum of contributions from independent diffusive mechanisms. The contribution of a particular microscopic mechanism can be written as the product of the diffusivity,  $D_i$ , and the concentration,  $C_i$ , of the relevant defect, *i.e.*,

$$D_{SD} = \sum_i f_i D_i C_i = \sum_i A_i \exp(-E_i/k_B T), \quad (1)$$

where the  $f_i$  are correlation factors of order unity,  $A_i$  is a prefactor and  $E_i$  is the activation energy for mechanism  $i$ . Various experimental studies have indicated that in silicon *two* mechanisms of self-diffusion are important, vacancy diffusion and self-interstitial diffusion. The prefactors and activation energies of the different mechanisms cannot be measured directly but their values can be deduced by combining various experimental data. Gösele *et al.* [7] have estimated that  $E_I = 4.84$  eV for interstitials and  $E_V = 4.03$  eV for vacancies, while Bracht *et al.* [8] have given  $E_I = 4.95$  eV and  $E_V = 4.14$  eV. The prefactor is predicted to be larger for vacancy diffusion than for self-interstitial diffusion and therefore vacancy diffusion dominates at low temperatures while self-interstitial diffusion dominates at higher temperatures, with a crossover at about 1100–1200 K. The situation is highly controversial when it comes to determining the individual values of  $D_I$  and  $C_I$ . Indeed, experimental data has been used to support values of the diffusivity of the silicon self-interstitial,  $D_I$ , which differ by ten orders of magnitude at the temperatures of around 800°C at which silicon is processed [9].

Self-diffusion in silicon is both technologically important and imperfectly understood, and our aim is to perform an accurate theoretical study which can help clarify the situation. In this paper we describe quantum Monte Carlo calculations of the energies of self-interstitial defects in silicon, and compare our results with those of density functional theory studies and with experiment.

## 2. THE VMC AND DMC METHODS

In the VMC method [1–3] expectation values are calculated as integrals over the configuration space, which are evaluated using standard Monte Carlo techniques. The variational energy  $E_V$  is given by

$$\begin{aligned} E_V &= \frac{\int \Phi_T(\mathbf{R}) \hat{H} \Phi_T(\mathbf{R}) d\mathbf{R}}{\int \Phi_T(\mathbf{R}) \Phi_T(\mathbf{R}) d\mathbf{R}} \\ &= \frac{\int \Phi_T(\mathbf{R})^2 (\Phi_T(\mathbf{R}))^{-1} \hat{H} \Phi_T(\mathbf{R}) d\mathbf{R}}{\int \Phi_T(\mathbf{R})^2 d\mathbf{R}}, \quad (2) \end{aligned}$$

where  $\Phi_T$  is the *trial wave function* (assumed real),  $\hat{H}$  is the Hamiltonian, and  $\mathbf{R}$  denotes the  $3N$ -dimensional vector of the electron positions. In the second line of Eq. (2) we have rewritten the numerator in terms of the *local energy*  $E_L = \Phi_T(\mathbf{R})^{-1} \hat{H} \Phi_T(\mathbf{R})$ . This step is very important for the statistical evaluation of the energy. The energy can now be evaluated as an integral over a probability distribution,  $P(\mathbf{R}) = \Phi_T(\mathbf{R})^2 / \int \Phi_T(\mathbf{R})^2 d\mathbf{R}$ . Such an integral can be performed using Monte Carlo techniques based on the Metropolis algorithm which generates a sequence of points  $\mathbf{R}_i$  which are asymptotically distributed according to the required probability distribution. The variational energy is then given by

$$E_V = \int P(\mathbf{R}) E_L(\mathbf{R}) d\mathbf{R} \approx \sum_i E_L(\mathbf{R}_i), \quad (3)$$

where the electron configuration vectors  $\mathbf{R}_i$  are distributed according to  $P(\mathbf{R})$ . The statistical error

in this estimate of  $E_V$  falls as the inverse of the square root of the number of sampling points.

The accuracy of the VMC method is limited by the quality of the trial wave function, and in practice one has to use highly sophisticated correlated wave-functions. Our trial wave functions are of the Slater-Jastrow type:

$$\Phi_T(\mathbf{R}) = D^\uparrow D^\downarrow \exp[J(\mathbf{R})], \quad (4)$$

where  $D^\uparrow$  and  $D^\downarrow$  are Slater determinates of up- and down-spin single particle orbitals and  $\exp[J(\mathbf{R})]$  is the *Jastrow factor*. In the calculations described here the single particle orbitals are obtained from local density approximation (LDA) calculations. We use Jastrow factors, which correlate the motion of pairs of electrons, although higher-body correlations can also be included. The Jastrow factor includes a number of parameters whose values may be varied to optimize the wave function. For each of the defect systems our trial wave function contains 64 variational parameters, while for the perfect crystal calculation there were 32 parameters. The optimal values were obtained by minimizing the variance of the energy [11, 10], which has been found to be more numerically stable than minimizing the energy itself [12].

The DMC method is a stochastic projector method for solving the imaginary-time Schrödinger equation,

$$-\frac{\partial \Psi(\mathbf{R}, t)}{\partial t} = (\hat{H} - E_S)\Psi(\mathbf{R}, t), \quad (5)$$

where  $t$  is a real variable measuring the progress in imaginary time, and  $E_S$  is an energy offset. This equation has the property that an initial starting state decays towards the ground state wave function (provided they have a non-zero overlap). The time evolution of Eq. (5) may be followed using a stochastic technique. The wave function  $\Psi(\mathbf{R}, t)$  is represented by an ensemble of  $3N$ -dimensional electron configuration vectors,  $\{\mathbf{R}_i\}$ , whose time evolution is governed by Eq. (5). The Green function of Eq. (5) shows

that the rules of the evolution are random diffusive jumps of the configuration vectors from the kinetic term and removal/addition of configuration vectors from the potential energy term which acts like a rate term.

Unfortunately this simple algorithm suffers from two serious problems. The first is that we have implicitly assumed that  $\Psi$  is a probability distribution, although its fermionic nature means that it must have positive and negative parts. To deal with the problem of the fermion antisymmetry we use the *fixed-node approximation*. The nodal surface of a wave function is the surface on which it is zero and across which it changes sign. If we force the time evolution of Eq. (5) to maintain a nodal surface consistent with the fermionic antisymmetry we can simulate the equation separately in each nodal pocket. The fixed nodal surface can be imposed by considering the evolution of the *mixed distribution*  $f = \Phi_T \Psi$ , where  $\Phi_T$  is the known trial wave function discussed above. The nodal surface of  $\Psi$  is constrained to be the same as that of  $\Phi_T$  and therefore  $f$  can be interpreted as a probability distribution. The fixed-node DMC method generates the distribution  $f = \Phi_T \Psi$ , where  $\Psi$  is the best (lowest energy) wave function with the same nodes as  $\Phi_T$ . The second problem is less fundamental but in practice very severe. The required rate of removing/adding configuration vectors diverges when the potential energy diverges, which occurs whenever two electrons or an electron and a nucleus are coincident, leading to poor statistical behaviour. Using the  $f$  distribution also introduces importance sampling which greatly reduces the statistical noise. Normally we demand that  $\Phi_T$  has the correct cusp-like behaviour when two electrons or an electron and a nucleus are coincident [13], which removes the divergence in the removal/addition process.

Using the fixed-node approximation means that we solve independently in different nodal pockets, and at first sight it appears that we have to solve the Schrödinger equation in every nodal pocket, which would be impossible in large systems.

However, the *tiling theorem* for fermion ground states [14, 15] asserts that all nodal pockets are in fact equivalent and therefore one only needs to solve the Schrödinger equation in one of them.

A simulation proceeds as follows. First we pick an ensemble of a few hundred configuration vectors chosen from the variational distribution,  $P(\mathbf{R})$ . This ensemble is evolved in imaginary time according to the rules of the importance sampled imaginary time Schrödinger equation, which involves biased diffusion and addition/subtraction steps. The bias in the diffusion is caused by the importance sampling which directs the sampling towards parts of configuration space where  $\Phi_T$  is large. After a period of equilibration the configuration vectors start to trace out the probability distribution  $f(\mathbf{R})/\int f(\mathbf{R})d\mathbf{R}$ . We can then start to accumulate averages, in particular the DMC energy,  $E_D$ , which is given by

$$E_D = \frac{\int f E_L d\mathbf{R}}{\int f d\mathbf{R}} \approx \sum_i E_L(\mathbf{R}_i), \quad (6)$$

where the configuration vectors are distributed according to  $f(\mathbf{R})/\int f(\mathbf{R})d\mathbf{R}$ . This energy expression would be exact if the nodal surface of  $\Phi_T$  was exact, and the fixed-node error is second order in the error in the nodal surface of  $\Phi_T$ . Fixed-node DMC calculations have now been performed for many systems. The accuracy of the fixed node approximation can be tested on small systems and normally leads to very satisfactory results [3].

This completes our brief introduction to VMC and DMC methods. We have left out many technical details of the calculations, and for more information on these aspects the reader is directed to Refs. [2, 3, 16].

### 3. STRUCTURES OF THE SELF-INTERSTITIAL DEFECTS

Experiments give no information about the structure of single-atom self-interstitial defects in

silicon. There have, however, been many theoretical studies of their structures. The most advanced of these have used the local density approximation (LDA) to density-functional theory to calculate the defect formation energies and energy barriers to diffusion [17–19]. The consensus view is that the split- $\langle 110 \rangle$ , hexagonal and tetrahedral defects (see Fig. 1) are low in energy. Other defects, such as the split- $\langle 100 \rangle$  and bond centred interstitials, are calculated to be much higher in energy and therefore we do not consider them here. We have also studied the saddle point of Pandey's concerted exchange mechanism [20] for self-diffusion, which involves the exchange of neighbouring atoms in the perfect lattice *via* a complicated 3-dimensional path which allows the atoms to avoid large energy barriers.

As a preliminary to our main calculations we performed a thorough study at the LDA level [21] using a plane wave basis set and a norm-conserving pseudopotential to represent the  $\text{Si}^{4+}$  ions. We performed LDA calculations for the fully relaxed defect structures using supercells for which the perfect crystalline structure contains 16, 54 and 128 atoms, respectively. The relaxed defect structures are almost independent of the size of the supercell, and are illustrated in Figure 1. One of the intriguing features of the various defect structures is the wide range of interatomic bonding they exhibit. The bond length for the perfect crystal is 2.35 Å. The nearest-neighbour defect bonds in the tetrahedral interstitial structure are significantly longer than the perfect crystal bonds at 2.44 Å, while at the saddle point of the concerted exchange the atoms are joined by a very short bond of length 2.15 Å. The variation in coordination number of atoms in the different defect structures varies from 3 (saddle point of concerted exchange) to 6 (hexagonal interstitial). This wide range of interatomic bonding poses a significant challenge to theoretical techniques.

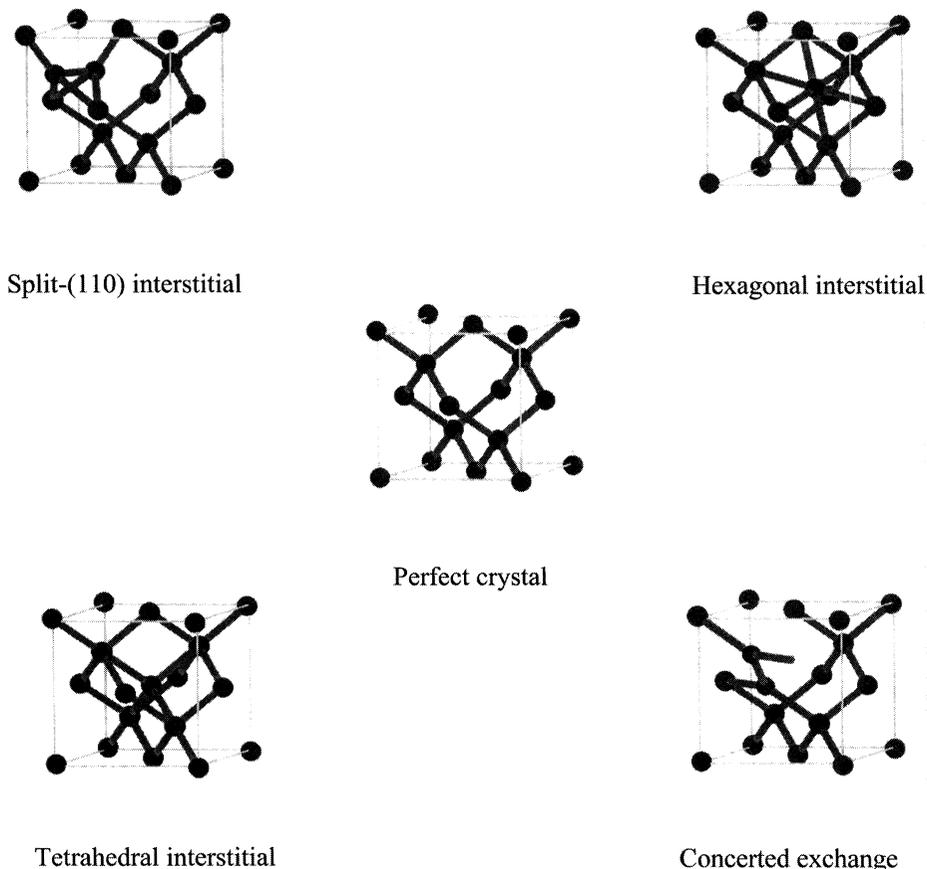


FIGURE 1 The split- $\langle 110 \rangle$ , hexagonal, and tetrahedral interstitial defects, the saddle point of the concerted-exchange mechanism, and the perfect silicon crystal.

#### 4. DEFECT FORMATION ENERGIES

The defect formation energies of the different structures calculated within the LDA are reported in Table I. We also show values for the defect energies calculated with the gradient-corrected PW91-GGA density functional [22]. From Table I we can see that the split- $\langle 110 \rangle$  and hexagonal interstitials are the lowest energy defects with DFT formation energies of about 3.3 eV (LDA) and 3.8 eV (PW91-GGA). We have also calculated the energy barriers to diffusive jumps between the low energy structures within DFT, finding a path for split- $\langle 110 \rangle$ -hexagonal diffusion with a barrier of

0.15 eV (LDA) and 0.20 eV (PW91-GGA), and a barrier for hexagonal-hexagonal diffusive jumps of 0.03 eV (LDA) and 0.18 eV (PW91-GGA). Our DFT results give the sum of the formation and migration energies for self-interstitials as 3.5 eV (LDA) and 4.0 eV (PW91-GGA), which are considerably smaller than the activation energy for self-interstitial diffusion deduced from experimental measurements of 4.84 eV [7] and 4.95 eV [8]. Although the LDA activation energy of 4.45 eV and the PW91-GGA activation energy of 4.80 eV for Pandey's concerted exchange are within the experimental range for self-diffusion of 4.1–5.1 eV [6], our results indicate that the activation

TABLE I LDA, PW91-GGA, and DMC formation energies in eV of the self-interstitial defects and the saddle point of the concerted-exchange mechanism

Defect	LDA	GGA	DMC
Split-110	3.31	3.84	4.96(28)
Hexagonal	3.31	3.80	4.82(28)
Tetrahedral	3.43	4.07	5.40(28)
Concerted Exchange	4.45	4.80	5.78(27)

energies for interstitial mediated self-diffusion are lower.

We conclude that the DFT results for self-interstitial diffusion in silicon do not afford a satisfactory explanation of the experimental temperature dependence of the self-diffusivity. Given the doubts about the accuracy of the DFT functionals and the technological importance of the problem, we have recalculated the energies of the defect structures using the DMC method using the 16 and 54 atom supercell-structures obtained from our LDA study. First we calculated the cohesive energy of silicon within the VMC and DMC methods by performing calculations for the atomic ground state and for the perfect solid using the 54-atom simulation cell. The resulting VMC and DMC cohesive energies of 4.48(1)eV and 4.63(2)eV per atom, respectively, are in excellent agreement with the experimental value of 4.62(8)eV per atom.

The results of our DMC calculations for the defect formation energies are shown in Table I. The clearest conclusion is that the DMC formation energies are roughly 1eV larger than the PW91-GGA values and 1.5eV larger than the LDA values. Within DMC the hexagonal interstitial has the lowest formation energy, and the split- $\langle 110 \rangle$  interstitial is slightly higher in energy, while the tetrahedral interstitial has a significantly higher energy. The saddle point of the concerted exchange has an energy which is too high to explain self-diffusion in silicon.

Since mapping the energy barriers to diffusion is computationally prohibitive within DMC we estimate the migration energy as follows. The tetrahedral interstitial is a saddle point of a possible diffusion path between neighbouring

hexagonal sites. The DMC formation energy of 5.4eV for the tetrahedral interstitial is therefore an upper bound to the formation+migration energy of the hexagonal interstitial. The true formation+migration energy is expected to be less than this, and we can use the following argument to obtain a crude estimate of it. Within the LDA we have found a diffusion path for the hexagonal interstitial with a barrier of only 25% of the tetrahedral-hexagonal energy difference. Applying the same percentage reduction to the DMC barrier gives an estimate of the formation+migration energy of the hexagonal interstitial of 5eV. This estimate is in good agreement with experimental values of the activation energy for self-interstitial diffusion of 4.84eV [7] and 4.95eV [8].

## 5. CONCLUSIONS

In conclusion we have given a brief introduction to the VMC and DMC methods and described their application to the problem of the energetics of self-interstitial defects in silicon. Our calculations demonstrate the importance of a proper treatment of electron correlation when calculating defect formation energies in silicon. The LDA and PW91-GGA functionals do not provide a satisfactory explanation of the experimental temperature dependence of the self-interstitial contribution to the self-diffusivity because they predict formation+migration energies which are too small. The larger defect formation energies found in our DMC calculations indicate a possible resolution of this problem. This may be an important step in improving our understanding of self-diffusion in silicon.

### Acknowledgments

We thank Paul Kent and Randy Hood for help with the calculations. Financial support was provided by the Engineering and Physical Sciences Research Council (UK), Hitachi Europe Ltd. and Hitachi Japan Ltd. The DMC calculations were performed on the Hitachi SR2201 at the University of Cambridge High Performance Computing Facility and at Hitachi's Central Research Laboratory.

### References

- [1] McMillan, W. L. (1965). *Phys. Rev.*, **138**, A442; Ceperley, D., Chester, G. V. and Kalos, M. H. (1977). *Phys. Rev. B*, **16**, 3081.
- [2] Ceperley, D., Chester, G. and Kalos, M. (1971). *Phys. Rev. B*, **16**, 3081; Hammond, B. L., Lester, W. A. Jr. and Reynolds, P. J., *Monte Carlo Methods in ab initio quantum Chemistry* (World Scientific, Singapore, 1994).
- [3] Hammond, B. L., Lester, W. A. Jr. and Reynolds, P. J., *Monte Carlo Methods in ab initio quantum Chemistry* (World Scientific, Singapore, 1994).
- [4] Leung, W.-K., Needs, R. J., Rajagopal, G., Itoh, S. and Ihara, S. (1999). *Phys. Rev. Lett.*, **83**, 2351.
- [5] Fahey, P. M., Griffin, P. B. and Plummer, J. D. (1989). *Rev. Mod. Phys.*, **61**, 289.
- [6] Frank, W., Gösele, U., Mehrer, H. and Seeger, A., In: *Diffusion in Crystalline Solids*, Edited by Murch, G. E. and Nowick, A. S. (Academic, Orlando, FL, 1985), p. 64.
- [7] Gösele, U., Plössl, A. and Tan, T. Y., In: *Process Physics and Modeling in Semiconductor Technology*, Edited by Srinivasan, G. R., Murthy, C. S. and Dunham, S. T. (Electrochemical Society, Pennington, NJ, 1996), p. 309.
- [8] Bracht, H., Haller, E. E. and Clark-Phelps, R. (1998). *Phys. Rev. Lett.*, **81**, 393.
- [9] Eaglesham, D. (1995). *Physics World*, **8**(11), 41.
- [10] Umrigar, C. J., Wilson, K. G. and Wilkins, J. W. (1988). *Phys. Rev. Lett.*, **60**, 1719.
- [11] Williamson, A. J., Kenny, S. D., Rajagopal, G., James, A. J., Needs, R. J., Fraser, L. M., Foulkes, W. M. C. and Maccallum, P. (1996). *Phys. Rev. B*, **53**, 9640.
- [12] Kent, P. R. C., Needs, R. J. and Rajagopal, G. (1999). *Phys. Rev. B*, **59**, 12344.
- [13] Kato, T. (1957). *Comm. Pure Appl. Math.*, **10**, 151.
- [14] Ceperley, D. (1991). *J. Stat. Phys.*, **63**, 1237.
- [15] Foulkes, W. M. C., Hood, R. Q. and Needs, R. J. (1999). *Phys. Rev. B*, **60**, 4558.
- [16] Foulkes, W. M. C., Mitas, L., Needs, R. J. and Rajagopal, G. (2001). *Rev. Mod. Phys.*, **73**, 33.
- [17] Bar-Yam, Y. and Joannopoulos, J. D. (1984). *Phys. Rev. B*, **30**, 1844.
- [18] Blöchl, P. E., Smargiassi, E., Car, R., Laks, D. B., Andreoni, W. and Pantelides, S. T. (1993). *Phys. Rev. Lett.*, **70**, 2435.
- [19] Needs, R. J. (1999). *J. Phys.: Condens. Matter*, **11**, 10437.
- [20] Pandey, K. C. (1986). *Phys. Rev. Lett.*, **57**, 2287; Kaxiras, E. and Pandey, K. C. (1993). *Phys. Rev. B*, **47**, 1659.
- [21] We used the LDA of Perdew, J. P. and Zunger, A. (1981). *Phys. Rev. B*, **23**, 5048 who parametrized the DMC data of Ceperley, D. M. and Alder, B. J. (1980). *Phys. Rev. Lett.*, **45**, 566.
- [22] Perdew, J. P., In: *Electronic Structure of Solids '91*, Edited by Ziesche, P. and Eschrig, H. (Akademie Verlag, Berlin, 1991), p. 11.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

