



**HAL**  
open science

# Substitutional carbon defects in silicon: A quantum mechanical characterization through the infrared and Raman spectra

Francesco Gentile, Alexander Platonenko, Khaled El-Kelany, Michel Rérat, Philippe d'Arco, Roberto Dovesi

## ► To cite this version:

Francesco Gentile, Alexander Platonenko, Khaled El-Kelany, Michel Rérat, Philippe d'Arco, et al.. Substitutional carbon defects in silicon: A quantum mechanical characterization through the infrared and Raman spectra. *Journal of Computational Chemistry*, 2020, 41 (17), pp.1638-1644. 10.1002/jcc.26206 . hal-02540365

**HAL Id: hal-02540365**

**<https://hal.science/hal-02540365v1>**

Submitted on 16 Apr 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Substitutional carbon defects in Silicon. A quantum mechanical characterization through the IR and Raman spectra.

Gentile F.S.\*; Platonenko A.†; El-Kelany K.E.‡; Rérat M.§; D’arco P. ¶; Dovesi R. ||

February 10, 2020

## Abstract

The Infrared and Raman spectra of eight substitutional carbon defects in silicon are computed at the quantum mechanical level by using a periodic supercell approach based on hybrid functionals, an *all electron* Gaussian type basis set and the CRYSTAL code. The single substitutional  $C_s$  case and its combination with a vacancy ( $C_sV$  and  $C_sSiV$ ) are considered first. The progressive saturation of the four bonds of a Si atom with C is then examined. The last set of defects consists of a chain of adjacent carbon atoms  $C_s^i$ , with  $i = 1$  to 3. The simple substitutional case,  $C_s$ , is the common first member of the three sets. All these defects show important, very characteristic features in their IR spectrum. One or two C related peaks dominate the spectra: at  $596\text{ cm}^{-1}$  for  $C_s$  (and  $C_sSiV$ , the second neighbor vacancy is not shifting the  $C_s$  peak), at  $705$  and  $716\text{ cm}^{-1}$  for  $C_sV$ , at  $537\text{ cm}^{-1}$  for  $C_s^2$  and  $C_s^3$  (with additional peaks at  $522$ ,  $655$  and  $689$  for the latter only), at  $607$  and  $624\text{ cm}^{-1}$ ,  $601$  and  $643\text{ cm}^{-1}$  and  $629\text{ cm}^{-1}$  for  $SiC_s^2$ ,  $SiC_s^3$ ,  $SiC_s^4$ , respectively. Comparison with experiment allows to attribute many observed peaks to one of the C substitutional defects. Observed peaks above  $720\text{ cm}^{-1}$  must be attributed to interstitial C or more complicated defects.

**Keywords:** Silicon, substitutional carbon defect, Ab-initio calculation, IR and Raman spectra. ■

---

\*Dipartimento di Chimica, Università di Torino and NIS (Nanostructured Interfaces and Surfaces) Centre, Via P. Giuria 5, 10125 Torino, Italy

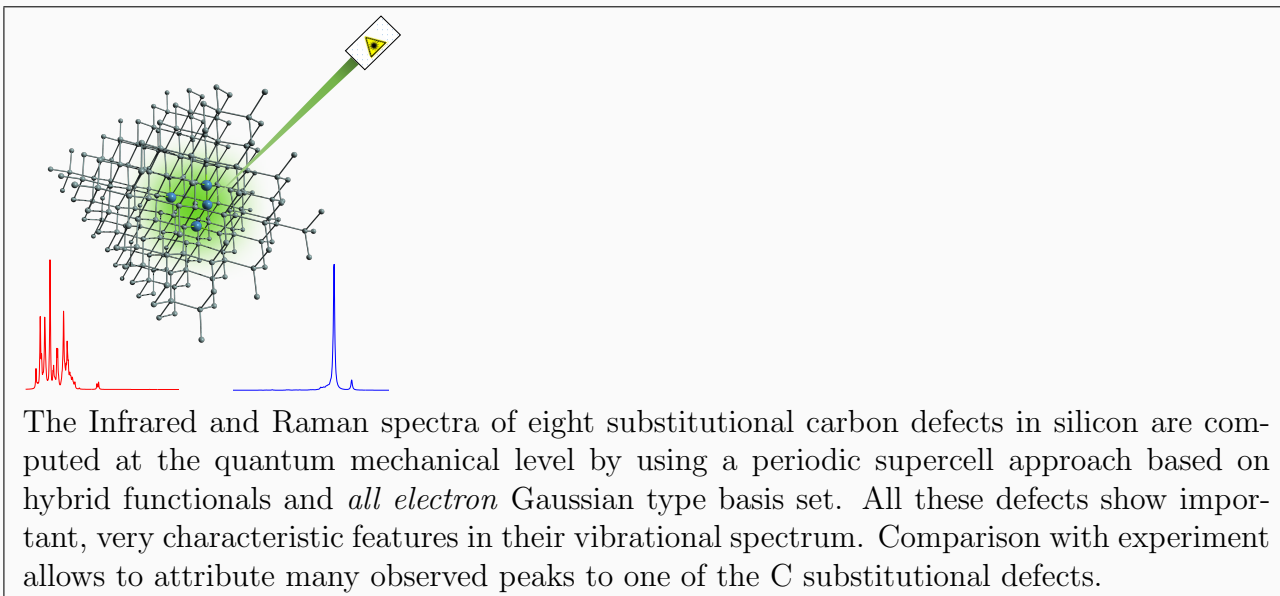
†Institute of Solid State Physics, University of Latvia, 8 Kengaraga street, LV1063, Riga, Latvia

‡Institute of Nanoscience and Nanotechnology, Kafrelshiekh University, 33516 Kafrelshiekh, Egypt

§Université de Pau et des Pays de l’Adour, E2S UPPA, CNRS, IPREM, 2 av. président P. Angot, Pau 64053, France

¶Sorbonne Universités, UPMC Paris 06, UMR 7193, Institut des Sciences de la Terre Paris (ISTeP), 4 place Jussieu, F-75005 Paris, France

||Dipartimento di Chimica, Università di Torino and NIS (Nanostructured Interfaces and Surfaces) Centre, Via P. Giuria 5, 10125 Torino, Italy



# INTRODUCTION

Electronic and mechanical properties of silicon are known to be significantly affected by the presence of impurities and defects. Carbon, along with oxygen, represents one of the most common impurity. Carbon is substitutionally dissolved as isovalent impurity and can interact with other point defects. It was found that implantation-generated excess of silicon self-interstitials after boron implantation is greatly reduced in the presence of carbon<sup>1</sup>. Carbon substitutional is considered immobile<sup>2</sup> and characterized by single vibrational mode at 608 cm<sup>-1</sup>, while carbon interstitial is highly mobile<sup>3</sup>. Carbon substitutional reacts with silicon self-interstitials, producing a new carbon interstitial defect. EPR measurements suggest that this defect has  $C_{2v}$  symmetry and a doublet electronic ground state.

In the present study we investigate eight defects in silicon, limiting the analysis to substitutional carbon. The aim is to explore whether the IR and/or Raman spectra provide enough information for the clear identification of the most important of these defects. As these defects are EPR inactive (they have a closed shell ground state, or the most stable state is a singlet, as for the two here considered that couple  $C_s$  and V, a vacancy), the vibrational analysis is the most general tool available for this characterization. The extension to defects involving interstitial carbon, possibly in combination with N or O will be the object of a forthcoming paper<sup>4</sup>.

The eight investigated defects can be grouped in three sets, the first one includes the simple substitutional  $C_s$  defect, and its *perturbed* cases  $C_sV$  and  $C_sSiV$ , in which  $C_s$  is a first or second neighbor of a vacancy. The second set includes 4 cases, to be indicated as  $SiC_s^i$ , with  $i=1$  to 4, in which the four neighbors of a Si atom are progressively substituted by C atoms. The third set includes three defects,  $C_s^i$ , with  $i=1$  to 3, in which a chain of adjacent carbon atoms has been taken into account.

The supercell scheme (cubic supercells containing 64 and 216 atoms will be considered), the very effective *hybrid* B3LYP<sup>5,6</sup> functional, an *all electron* Gaussian type basis set and the CRYSTAL code have been used. As outlined previously<sup>7-12</sup>, the B3LYP hybrid functional, with 20% non-local exact exchange, leads to a proper description of (electronic) spin polar-

ization, and accurate vibrational spectra.

The paper is structured as follows. Section METHODOLOGY contains the details of the computational procedures and conditions. Section RESULTS describes the geometry, the charge and spin (when applies) density, and the vibrational features of the eight defects. In Section DISCUSSION the results are discussed and compared with experiments. A few conclusions are drawn in section CONCLUSIONS. A Supplementary Material Section contains some of the technical details concerning the vibrational part, and shows the Raman spectra of the eight defects.

## METHODOLOGY

Calculations have been performed by use of the B3LYP global hybrid functional,<sup>5,6</sup> as implemented in the CRYSTAL17 rogram.<sup>13</sup> An *all electron* Gaussian type basis set (derived from Pople’s standard 6-31G<sup>14</sup>) as described in D’Arco *et al.*<sup>15</sup> has been adopted for silicon: the contraction is 6-5111G\* (22 atomic orbitals, AOs), where the first shell is of *s* type, followed by 4 *sp* and one d shell; the exponents of the two outermost *sp* shells have been set to 0.45 Bohr<sup>-2</sup> and 0.15 Bohr<sup>-2</sup>, respectively. For the carbon atom, the 6-31G\*<sup>16</sup> contraction has been adopted without modifications (14 AOs both).

The truncation of the Coulomb and exchange infinite lattice series is controlled by five thresholds  $T_i$  (see CRYSTAL manual, Ref. 17, for more details), which have been set to 8 ( $T_1$ - $T_4$ ) and 16 ( $T_5$ ). The convergence threshold on energy for the self-consistent-field (SCF) procedure has been set to  $10^{-8}$  and  $10^{-11}$   $E_h$  for the structural and frequency calculations, respectively.

The DFT exchange-correlation contribution to the Fock matrix has been evaluated by numerical integration over the unit cell volume. Radial and angular points for the integration grid are generated through Gauss-Legendre radial quadrature and Lebedev two-dimensional angular point distributions. The default pruned grid with 74 radial and 974 angular points has been used.

A periodic supercell approach is used to simulate different defect concentrations. Cubic

super-cells containing 64 and 216 atoms have been considered (to be referred to in the following as  $S_n$ , with  $n= 64$  or  $216$ ). They correspond to the expansion of the bulk silicon conventional cell, containing 8 atoms, by a factor 2 and 3 along the three lattice parameters. The defects reduce the symmetry; the deformation of the cell from the cubic shape is however negligible. Overall, the size of the basis set (number of AOs) is (about) 1400 in  $S_{64}$  and (about) 4744 in  $S_{216}$ . Reciprocal space has been sampled by using a regular sub-lattice with a shrinking factor 4 for  $S_{64}$  and 2 for  $S_{216}$ ; this corresponds to a number of  $\mathbf{k}$ -points ranging from 4 to 36, according to the point symmetry of the defective system.

The experimental Raman spectrum of pristine silicon<sup>18</sup> is characterized by a single, three fold degenerate, peak at  $521\text{ cm}^{-1}$ . The calculated peak with the present functional and basis set is at  $533\text{ cm}^{-1}$ ,  $12\text{ cm}^{-1}$  above. So any observed IR signal or any Raman peak other than the one at  $521\text{ cm}^{-1}$  must be attributed to defects.

As the mass of carbon is less than half the one of silicon (12.01 vs 28.09 u), we can anticipate that many of the modes to which C contributes will appear well above the phonon band of pristine silicon, that has a calculated higher limit at  $533\text{ cm}^{-1}$ . Besides that, the C-Si bonds are polar (typical electronegativity values are 1.9 for Si and 2.6 for C); as a consequence, the IR *defect peaks* are expected to be very intense, and to dominate the spectrum.

The identification of the nature of the modes (which atom is contributing? what kind of movement is described by the mode?) can easily be performed through two tools implemented in the CRYSTAL code:

a) the isotopic shift (see the note following eq. 1 in the Supplementary Material Section ); at negligible cost the mass of a subset of atoms can be altered arbitrarily. Note that at variance with respect to experiment, i) only very particular atoms of the cell can be involved in the isotopic substitution; ii) it is not necessary that the attributed mass corresponds to an existing isotope. Very large masses can be attributed to atoms in order to magnify the mass effect.

b) the graphical animation of the modes. This permits to have a direct and clear evidence of the atoms and groups providing important contributions to the eigenvectors; it is available at [www.crystal.unito.it](http://www.crystal.unito.it) in the *Defects in Silicon* section.

## RESULTS

In this section we consider eight defects containing substitutional carbon, organized in three series.

The first series contains the isolated substitutional carbon,  $C_s$ , and the combination of  $C_s$  with a vacancy  $V$ :  $C_sSiV$ , when  $C$  is a second neighbor of  $V$ , and  $C_sV$ , when  $C_s$  is sitting on the border of  $V$ . The ground state of  $C_s$  is a singlet closed shell, whereas in the two other cases the four first neighbors of the vacancy carry a uncoupled electron, so that a quintuplet  $q$  ( $S_z=2$ ), a triplet  $t$  ( $S_z=1$ ) and a singlet  $s$  ( $S_z=0$ ) spin states are possible. The singlet is more stable than the triplet by 0.10 ( $C_sSiV$ ) and 0.72 eV ( $C_sV$ ), respectively. The  $q$  state is less stable than the triplet (see Table 2).

In Figure 1, top, these three defects are schematically represented, with atom-atom distances  $R$  and Mulliken charges  $q$  and bond populations  $b$ . The electronegativity of  $C$  is larger than the  $Si$  one (2.6 vs 1.9), so that the Mulliken net charge  $q$  of  $C$  is  $-0.29 |e|$  in  $C_s$ ; it increases (in absolute value) to  $-0.44$  and  $-0.55 |e|$  in the two defects including the vacancy. The  $R$  distances, that are 1.56 Å in pristine diamond and 2.57 Å in silicon, are in between these values for the various  $C$ - $Si$  distances. The bond populations  $b$  are intermediate between the ones of diamond and silicon.

The IR spectrum of  $C_s$  is very simple (see Figure 3, left). It is dominated by a single peak at  $596 \text{ cm}^{-1}$ , whose intensity is 2015  $\text{km/mol}$ . In  $C_sSiV$  the peak splits in three, and blue shifts by a few  $\text{cm}^{-1}$  (600, 601 and  $617 \text{ cm}^{-1}$ ); the sum of the intensities of the three peaks is about the same as the one of  $C_s$ . When carbon is on the border of the vacancy, the shift is larger (more than  $100 \text{ cm}^{-1}$ ). The lower symmetry splits the single peak of  $C_s$  in two peaks at close wavenumbers, 705 and  $716 \text{ cm}^{-1}$ , with a much smaller intensity.

The second series includes  $C_s$ ,  $C_s^2$  and  $C_s^3$ , a chain of 1 to 3 adjacent substitutional carbon atoms. The ground state is always a closed shell; charges and distances are shown in Figure 1, second row. The carbon net charge is as high as  $-0.55$  and  $-0.58 |e|$  in  $C_s^2$ , and in  $C_s^3$  (on the two terminal atoms; for the central one it is  $-0.31 |e|$ ). The  $C$ - $C$  distances are much

larger than in diamond (1.81 and 1.97 Å in the double and triple substitutional case), due to the rigidity of the bulk Si lattice.

As regards the IR spectrum, in  $C_s^2$  and  $C_s^3$  the 596  $\text{cm}^{-1}$   $C_s$  peak red shifts by 60  $\text{cm}^{-1}$  to 537  $\text{cm}^{-1}$ . In  $C_s^3$ , three additional peaks of lower intensity appear at 585 (11  $\text{cm}^{-1}$  below the  $C_s$  peak), 655 and 689  $\text{cm}^{-1}$ .

The former corresponds to a C-C stretching, the latter to the C-Si stretching of the central carbon atom, as the graphical animation of the modes confirms. The 585  $\text{cm}^{-1}$  mode corresponds to C-Si stretching of two outer carbon atoms. The C-Si peak is absent in  $C_s^2$  for symmetry reasons. It appears on the contrary in the Raman spectrum, see the Supplementary Material, Figure S1). The mode at 651  $\text{cm}^{-1}$  corresponds to the asymmetric C-Si stretching.

The third series includes  $\text{SiC}_s^1 \equiv C_s$ ,  $\text{SiC}_s^2$ ,  $\text{SiC}_s^3$  and  $\text{SiC}_s^4$  (second row in Figure 1), corresponding to substitute one, two, three and four neighbors of a central silicon atom. The carbon net charge, that in  $C_s$  is -0.290  $|e|$ , and is more than compensated by the four Si atoms (+0.110  $|e|$ ), jumps to -0.824, -0.825 and -0.896  $|e|$  in the three other cases, whereas the central Si atom has a net charge of +0.475, +0.687 and +0.896  $|e|$ . The net charge of the "external" Si atoms is constant at about +0.240  $|e|$  for the three cases. The Si-C bond population reduces from +0.34  $|e|$  for  $C_s$  to about +0.26  $|e|$  for the three other cases, as a consequence of the increased distance between the central Si atom and its neighbors from 1.82 to 2.11, 2.16 and 2.19 Å along the series.

The IR spectrum of the three  $\text{SiC}_s^x$  defects, with  $x = 2$  to 4, is shown in Figure 3, right column. The peak at 596  $\text{cm}^{-1}$  of  $C_s$  splits in two peaks at 607 and 624  $\text{cm}^{-1}$  ( $\text{SiC}_s^2$ ), and at 601 and 643  $\text{cm}^{-1}$  ( $\text{SiC}_s^3$ ), by effect of the reduced symmetry ( $C_{2v}$  and  $C_{3v}$ , respectively). In  $\text{SiC}_s^4$  a single peak (symmetry is  $T_d$ ) at 629  $\text{cm}^{-1}$  dominates the spectrum, intermediate between the two peaks of  $\text{SiC}_s^2$  and  $\text{SiC}_s^3$ . The isotopic substitution  $^{12}\text{C} \rightarrow ^{13}\text{C}$  red shifts these peaks by 18-20  $\text{cm}^{-1}$ , confirming that they are related to carbon. Less intense peaks in the range 520-548  $\text{cm}^{-1}$  appear in the 3 cases. Their isotopic shift is only 3-5  $\text{cm}^{-1}$  large, indicating that they are due to the perturbation of the Si lattice around the defect, rather than being directly connected to C.



The defect formation energies  $E_f$  were calculated according to the following equation:

$$E_f = (E_D - n * E_C) - (E_P - m * E_{Si}) \quad (1)$$

where  $E_D$  and  $E_P$  are the total energies of the defective ( $D$ ) and pristine ( $P$ ) silicon supercells ( $S_{64}$  or  $S_{216}$ ),  $n$  and  $m$  are the numbers of carbon atoms added to, or silicon atoms removed from, the defective supercells. Formation energies for all studied defects are reported in Table 2.  $E_{Si}$  is the total energy of the silicon atom in the pristine silicon structure (that is, half the electronic total energy of the primitive cell, that with the present basis set and functional is  $-289.46807 E_h$  per atom). As regards the definition of  $E_C$ , we have followed two different approaches, denoted by superscripts 1 and 2. In the first one ( $E_f^1$ ),  $E_C$  is obtained in a similar way from pristine diamond ( $-38.08851046 E_h$  per atom). In the second definition,  $E_f^2$  is the energy of an isolated carbon atom. The difference between the two is obviously nothing else than the binding energy of the diamond crystal.

$E_f^1$  for  $C_s$  is 1.5 eV. The one of  $C_s^2$  is not twice the one of  $C_s$ , but about three times, and three more eV are necessary to create  $C_s^3$ . This discontinuity is confirmed by a parallel discontinuity in the  $C_sSi$  distances, and  $C_s$  net charges.

If we compare the two systems with two or three C atoms, the table shows that localizing the perturbation around a single Si atom is strongly favoured: 3.11 ( $SiC_s^2$ ) vs 4.57 ( $C_s^2$ ) eV, and 4.84 ( $SiC_s^3$ ) vs 7.71 ( $C_s^3$ ) eV. Looking at the  $SiC_s^i$  series, the progressive insertion of one, two, three and four C atoms requires 1.5, 1.61, 1.73 and 1.68 eV. We considered also the formation energy of C-Si-Si-C, when the two carbon atoms are linked to two first neighbors silicon atoms. The  $E_f^1$  formation energy is 3.88 eV, much lower than the  $C_s^2$  one (4.57 eV) but much higher than the  $SiC_s^2$  one (3.11 eV). It is clear that if  $n$  carbon atoms are present in bulk silicon, they will tend to aggregate.

The formation energy of a single silicon vacancy obtained from our model is about 4.7 eV (the experimental estimated value is 4.0 eV<sup>19</sup>).  $E_f^1$  for  $C_sSiV$  is 6.52 eV, slightly larger (0.32 eV) than the sum of the formation energy of the two defects (4.7 + 1.5 = 6.2 eV).  $E_f^1$  of  $C_sV$  is, on the contrary, exactly equal to the sum of  $E_f^1$  for V and  $C_s$ .

## DISCUSSION

Associating an experimental IR, Raman or EPR signal to a specific defect is challenging and, very often, controversial. In this section we try to connect the microscopic structure of defects to experiments comparing calculated and measured IR spectra.

In 1965 Newman<sup>2</sup> published the IR spectrum of carbon doped monocrystalline silicon. The spectra collected at 290 K were characterized by a single peak at 608 cm<sup>-1</sup>. For a sample containing both <sup>12</sup>C and an increased percentage of <sup>13</sup>C, two peaks at 604 and 586 cm<sup>-1</sup> were observed; the second one was attributed to <sup>13</sup>C, with an isotopic shift of 18 cm<sup>-1</sup>.

The dominant peak in our simulated spectrum of C<sub>s</sub> is at 596 cm<sup>-1</sup>, that downshifts by 18 cm<sup>-1</sup> to 578 cm<sup>-1</sup> when <sup>13</sup>C is used instead of <sup>12</sup>C (see Table 1). The present scheme (B3LYP plus the basis set described in the computational section) underestimates the dominant Raman peak of perfect bulk silicon by 12 cm<sup>-1</sup> (from 533 to 521 cm<sup>-1</sup>). It is reasonable to suppose that a similar underestimation is affecting the spectra of the defective systems. Taking into account this red shift, it can be concluded that the peaks observed by Newman can be attributed to a single carbon substitutional defect, C<sub>s</sub>.

In the same experiment two more modes were observed when collecting spectra at 63 K. For <sup>12</sup>C plus <sup>13</sup>C doped samples, absorption bands at 1217 and 1175 cm<sup>-1</sup> with much smaller intensities were attributed to possible second harmonics of the local mode vibrations. We are collecting some evidences that interstitial carbon defects could be, alternatively, responsible for these peaks<sup>4</sup>.

In 1970 Bean *et al.*<sup>20</sup> reported a large list of modes that could be associated to oxygen and carbon defects or their complexes. For samples containing only carbon, four low intensity modes in the range from 528 cm<sup>-1</sup> to 552 cm<sup>-1</sup> were reported, without attribution to specific local structures.

Our calculated IR spectrum of the C<sub>s</sub><sup>2</sup> defect (two vicinal substitutional carbon atoms, see figure 4), is dominated by a single peak at 537 cm<sup>-1</sup>, well within the range of the peaks reported by Bean *et al.* Other C substitutional defects show peaks in the 528-552 cm<sup>-1</sup> range: C<sub>s</sub><sup>3</sup>, with the two dominant peaks at 537 (as for C<sub>s</sub><sup>2</sup>) and 527 cm<sup>-1</sup>; SiC<sub>s</sub><sup>3</sup> and SiC<sub>s</sub><sup>4</sup> present low intensity peaks in the range 522-532 cm<sup>-1</sup>. Note however that in the case of

$\text{SiC}_s^3$  and  $\text{SiC}_s^4$  also the dominant peaks at 601 and 643  $\text{cm}^{-1}$  for the first, and at 629  $\text{cm}^{-1}$  for the second, should appear in the experimental spectrum. So the  $\text{C}_s^2$  and  $\text{C}_s^3$  defects are the best candidates for the attribution of these experimental peaks.

Bean *et al.*<sup>20</sup> reports also about a broad band in the range 704-735  $\text{cm}^{-1}$ . The authors suggest that this band could be related to C-V centers. Inspection of figures 3 and 4 shows that all C substitutional defects present peaks below 690  $\text{cm}^{-1}$ , the exception being  $\text{C}_s\text{V}$ , a carbon first neighbor of a vacancy, that in its singlet state (the most stable), has two dominant peaks at 705 and 716  $\text{cm}^{-1}$ . We can then confirm the attribution by Bean *et al.*

## CONCLUSIONS

The IR and Raman spectra, as well as the local geometry, the charge and spin (when applies) distribution of eight substitutional carbon defects are produced. The Raman spectrum is in all cases dominated by a peak at 530  $\text{cm}^{-1}$ , corresponding to the only peak of pristine bulk silicon. The landscape of the IR spectra is much more varied, and allows in most of the cases a clear identification of the specific defect.  $\text{C}_s$  has a single very intense peak at 596  $\text{cm}^{-1}$ , whereas  $\text{C}_s\text{V}$  is the only one with peaks above 700  $\text{cm}^{-1}$  (705 and 716  $\text{cm}^{-1}$ ).

With respect to  $\text{C}_s$ , the  $\text{SiC}_s^i$  defects blue shift to 607 and 624  $\text{cm}^{-1}$ , and red shift to 548  $\text{cm}^{-1}$  ( $\text{SiC}_s^2$ ), blue shift to 601 and 643  $\text{cm}^{-1}$  ( $\text{SiC}_s^3$ ), or to 629  $\text{cm}^{-1}$  ( $\text{SiC}_s^4$ ), whereas  $\text{C}_s^2$ ) and  $\text{C}_s^3$ ) red shift to 523-533  $\text{cm}^{-1}$ . The IR spectrum is then able to characterize all these defects.

The most relevant experimental peaks have been attributed to some of these defects.

Peaks above 720  $\text{cm}^{-1}$  are most probably attributable to interstitial C defects that will be the object of a forthcoming investigation.

The present study confirms that modern, high quality quantum-mechanical computer codes are extremely useful tools for the analysis of the properties of defects in crystalline compounds. They can be considered as standard tools, that complement in many cases, and substitute in others, experimental investigations, when the latter are difficult, or very expensive, or unable to discriminate between different hypotheses.

## **ACKNOWLEDGMENTS**

Access to the HPC resources of CINES/IDRIS/TGCC obtained thanks to the grant 2018-[A0050810537] (Ph. D'Arco) made by GENCI is warmly acknowledged.

((Additional Supporting Information may be found in the online version of this article.))

## References

1. P. Stolk, H.-J. Gossmann, D. Eaglesham, D. Jacobson, C. Rafferty, G. Gilmer, M. Jaraiz, J. Poate, H. Luftman, and T. Haynes, *Journal of Applied Physics* **81**, 6031 (1997).
2. R. Newman and J. Willis, *Journal of physics and chemistry of solids* **26**, 373 (1965).
3. R. a. Pinacho, P. Castrillo, M. Jaraiz, I. Martin-Bragado, J. Barbolla, H.-J. Gossmann, G.-H. Gilmer, and J.-L. Benton, *Journal of Applied Physics* **92**, 1582 (2002).
4. F. S. Gentile, A. Platonenko, and R. Dovesi, in preparation (2020).
5. A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
6. C. Lee, W. Yang, and R. Parr, *Phys. Rev. B* **37**, 785 (1988).
7. L. Bjaalie, A. Janotti, K. Krishnaswamy, and C. G. Van de Walle, *Phys. Rev. B* **93**, 115316 (2016).
8. M. Gerosa, C. Di Valentin, C. E. Bottani, G. Onida, and G. Pacchioni, *J. Chem. Phys.* **143**, 111103 (2015).
9. M. Gerosa, C. E. Bottani, L. Caramella, G. Onida, C. Di Valentin, and G. Pacchioni, *J. Chem. Phys.* **143**, 134702 (2015).
10. I. d. P. R. Moreira and R. Dovesi, *Int. J. Quantum Chem.* **99**, 805 (2004).
11. D. Munoz, N. M. Harrison, and F. Illas, *Phys. Rev. B* **69**, 085115 1 (2004).
12. J. C. Wojdeł, I. de P. R. de Moreira, S. T. Bromley, and F. Illas, *J. Chem. Phys.* **128**, 044713 1 (2008).
13. R. Dovesi, R. Orlando, A. Erba, C. M. Zicovich-Wilson, B. Civalleri, S. Casassa, L. Maschio, M. Ferrabone, M. D. L. Pierre, P. D'Arco, et al., *Int. J. Quantum Chem.* **114**, 1287 (2014).
14. M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, and J. A. Pople, *The Journal of Chemical Physics* **77**, 3654 (1982).

15. P. D'Arco, G. Sandrone, R. Dovesi, R. Orlando, and V. Saunders, *Physics and Chemistry of Minerals* **20**, 407 (1993).
16. M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, and J. A. Pople, *J. Chem. Phys.* **77**, 3654 (1982).
17. R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. J. Bush, et al., *CRYSTAL 2014 User's Manual*, (University of Torino, Torino, 2013).
18. J. H. Parker, D. W. Feldman, and M. Ashkin, *Phys. Rev.* **155**, 712 (1967), URL <https://link.aps.org/doi/10.1103/PhysRev.155.712>.
19. N. Fukata, A. Kasuya, and M. Suezawa, *Japanese Journal of Applied Physics* **40**, L854 (2001).
20. A. Bean, R. Newman, and R. Smith, *Journal of Physics and Chemistry of Solids* **31**, 739 (1970).

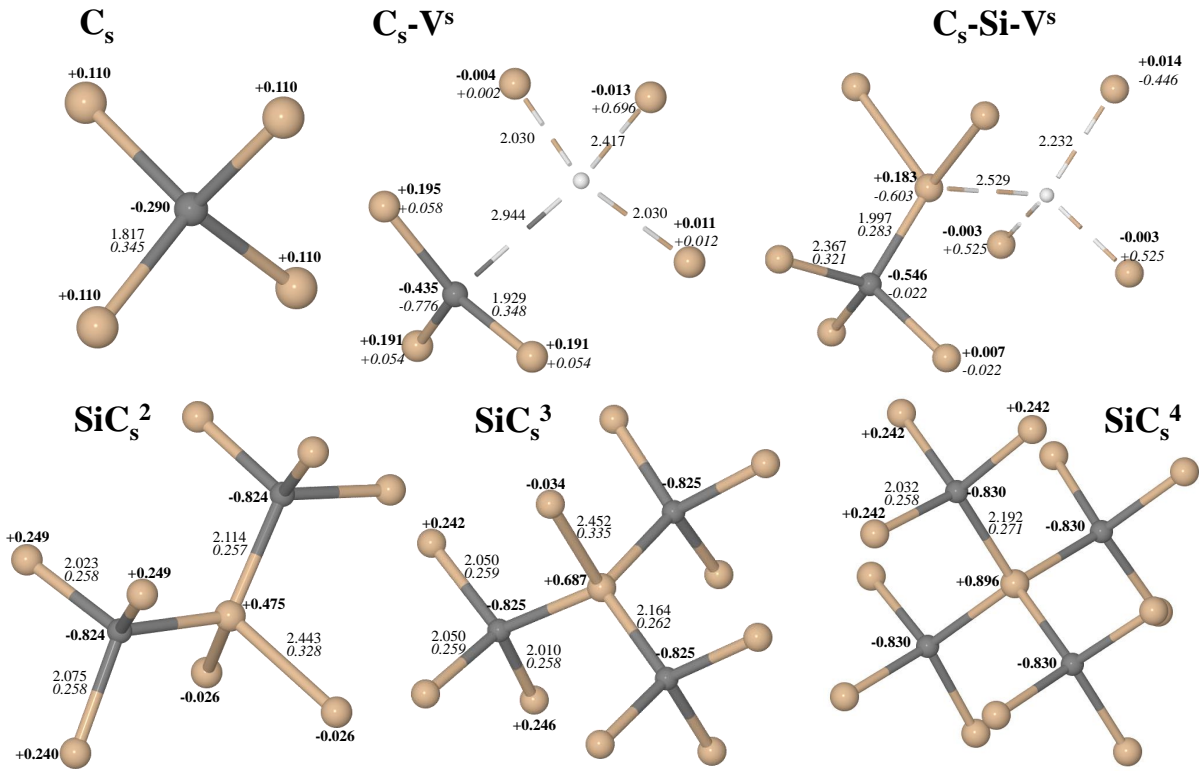


Figure 1: The local cluster around the substitutional carbon and vacancy defects here considered. Bond lengths (in Å), Mulliken net charges (in  $|e|$  and ***bold***) and bond populations (in  $|e|$  and *italic*) are reported for nearest non equivalent atoms. Superscript  $s$  indicates an open shell singlet. Carbon atoms are in grey, the vacancy position is identified by the white circle.

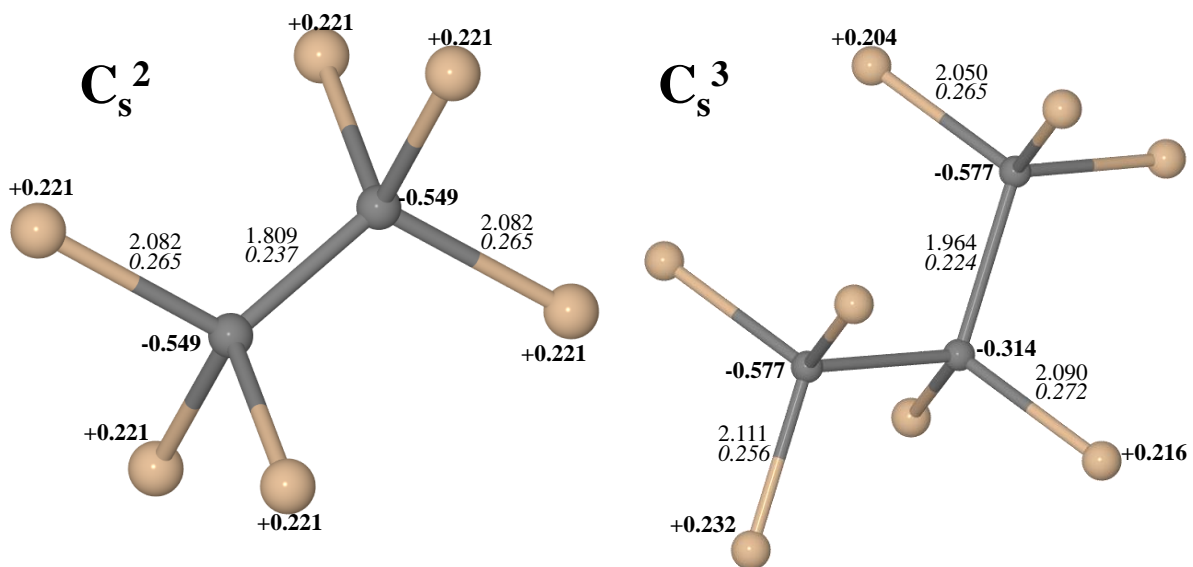


Figure 2: The local cluster around the  $C_s^x$  defects here considered. Bond lengths (in Å), Mulliken net charges (in  $|e|$  and **bold**) and bond populations (in  $|e|$  and *italic*) are reported for nearest non equivalent atoms. Carbon atoms are in grey.



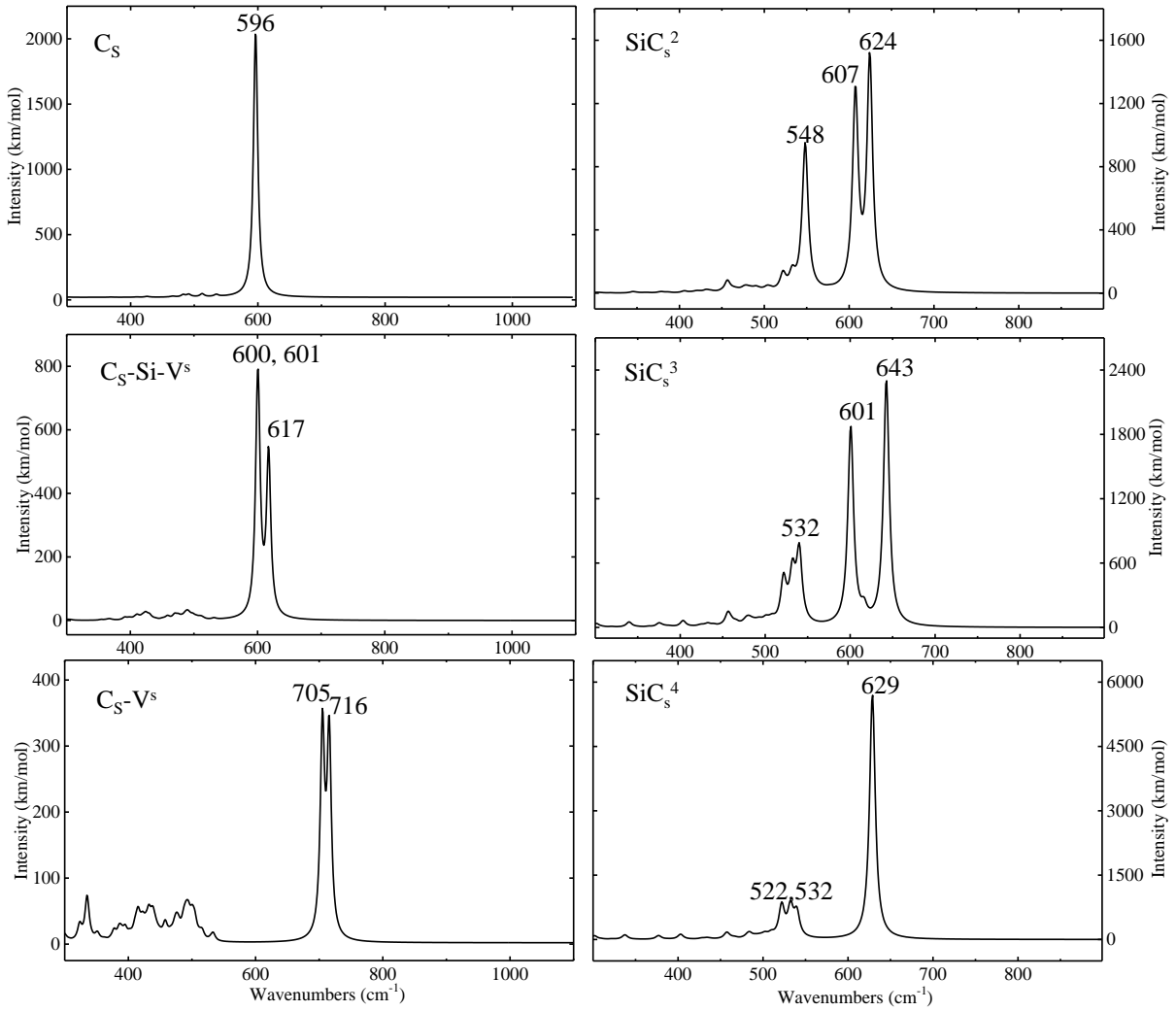


Figure 3: Simulated IR spectra of carbon substitutional defects.  $s$  superscript correspond to the singlet solution. In all other cases the solution is a closed shell.

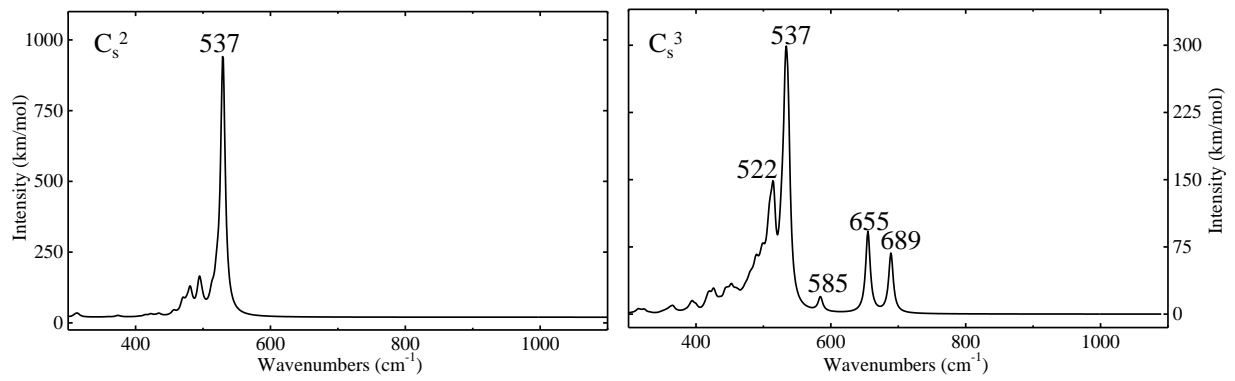


Figure 4: Simulated IR spectra of C<sub>s</sub><sup>x</sup> carbon substitutional defects.

| Defect                       | Sym             | $\nu(^{12}\text{C}) \rightarrow \nu(^{13}\text{C}), \delta; (\text{intensity})$   |
|------------------------------|-----------------|---|
| $\text{C}_s$                 | $\text{T}_d$    | 596→578, <b>18</b> (2015)   |
| $(\text{C}_s\text{-Si-V})^s$ | $\text{C}_s$    | 617→598, <b>19</b> (791); 601→583, <b>18</b> (632); 600→582, <b>18</b> (615)  |
| $(\text{C}_s\text{-V})^s$    | $\text{C}_s$    | 716→693, <b>23</b> (344); 705→683, <b>22</b> (355)  |
| $\text{C}_s^2$               | $\text{C}_{3v}$ | 651→628, <b>23</b> (R); 544→543, <b>1</b> (R); 537→537 <b>0</b> (941)   |
| $\text{C}_s^3$               | $\text{C}_{2v}$ | 689→664, <b>25</b> (144); 655→631, <b>24</b> (196); 612→592, <b>20</b> (R); 585→567, <b>8</b> (25); 537→536, <b>1</b> (299) |
| $\text{SiC}_s^2$             | $\text{C}_{2v}$ | 624→605, <b>19</b> (1522); 612→592, <b>20</b> (R); 1294, 548→540, <b>8</b> (943)  |
| $\text{SiC}_s^3$             | $\text{C}_{3v}$ | 642→623, <b>19</b> (2298); 601→582, <b>19</b> (1850); 540→535, <b>5</b> (653)   |
| $\text{SiC}_s^4$             | $\text{T}_d$    | 629→609, <b>20</b> (5692); 540→537, <b>3</b> (559)  |

Table 1: Wavenumber shift  $\delta$  (in  $\text{cm}^{-1}$  due to the  $^{12}\text{C} \rightarrow ^{13}\text{C}$  isotopic substitution) of the most intense IR modes. The absolute intensities (in  $\text{km}/\text{mol}$ ) are also reported. A semicolon separates the information concerning one mode from the one of another mode.

| Method                    | $E_D$<br>(Ha) | $E_f^1$ (eV) |      |      | $E_f^2$ (eV) |       |       |
|---------------------------|---------------|--------------|------|------|--------------|-------|-------|
|                           |               | s            | t    | Q    | s            | t     | Q     |
| $\text{C}_s$              | -18274.52383  | 1.50         |      |      | -6.55        |       |       |
| $\text{C}_s\text{-Si-V}$  | -17984.83849  | 6.52         | 6.61 | 7.42 | -1.54        | -1.44 | -0.64 |
| $\text{C}_s\text{-V}$     | -17984.85291  | 6.19         | 6.92 | 7.02 | -1.86        | -1.14 | -1.03 |
| $\text{C}_s^2$            | -18023.03141  | 4.57         |      |      | -11.53       |       |       |
| $\text{C}_s^3(S_{216})$   | -61770.68856  | 7.71         |      |      | -16.45       |       |       |
| $\text{SiC}_s^2(S_{216})$ | -62022.23700  | 3.11         |      |      | -12.99       |       |       |
| $\text{SiC}_s^3(S_{216})$ | -61770.79382  | 4.84         |      |      | -19.31       |       |       |
| $\text{SiC}_s^4(S_{216})$ | -61519.35278  | 6.52         |      |      | -25.69       |       |       |

Table 2: Total ( $E_D$ , in  $E_h$ ) and formation  $E_f^1$  (in eV) energy of the substitutional defects in silicon (see equation 1), computed with reference to pristine diamond and silicon bulk, and to the isolated carbon atom. The  $S_{64}$  supercell has been used for the first four defects,  $S_{216}$  for the more extended defects, as indicated.