Decoupled DFT-$\frac{1}{2}$ method for defect excitation energies

The decoupled DFT-$\frac{1}{2}$ method for defect excitation energies

Joshua Claes and Bart Partoens
CMT, Physics Department, University of Antwerp. Groenenenborgerlaan 171, 2020 Belgium

Dirk Lamoen
EMAT, Physics Department, University of Antwerp. Groenenenborgerlaan 171, 2020 Belgium
(Dated: August 30, 2023)

The DFT-$\frac{1}{2}$ method is a band gap correction with GW precision at a DFT computational cost. The method was also extended to correct the gap between defect levels, allowing for the calculation of optical transitions. However, this method fails when the atomic character of the occupied and unoccupied defect levels are similar as we illustrate by two examples, the tetrahedral hydrogen interstitial and the negatively charged vacancy in diamond. We solve this problem by decoupling the effect of the occupied and unoccupied defect levels and call this the decoupled DFT-$\frac{1}{2}$ method for defects.

I. INTRODUCTION

Creating point defects in solids allows one to manipulate the properties of that solid for example by doping the material to obtain a P or N type semiconductor. The defects themselves can also be interesting subsystems in the context of quantum technologies, where defects like the NV center [1, 2] have the potential to be used as a quantum sensor or as qubits, the building block for the quantum computer [3, 4]. The primary approach to simulate these defects is density functional theory (DFT). However, DFT is known to underestimate the band gap of semiconductors for the LDA or GGA exchange correlation functionals. This is due to the local nature of these approximate exchange correlation functionals, which neglects the exchange correlation discontinuity at the band gap [5–7]. When the band gap is a property of interest, as is the case with defects in solids, one has to rely on more advanced methods such as meta-GGA functionals like SCAN [8] or hybrid functionals like HSE06 [9, 10] or GW.

In 2008 Ferreira et al. [11] introduced a new method, the DFT-$\frac{1}{2}$ method, which rectifies the lack of self-energy of the band gap. This is achieved by adding a self-energy potential to the pseudopotential in such a way that the self-energy is added to the band gap. The DFT-$\frac{1}{2}$ method was later expanded upon by Lucatto et al. [12] to work for defect levels. However, when these defect levels have a similar orbital character, the DFT-$\frac{1}{2}$ method will fail. Prior to this the DFT-$\frac{1}{2}$ method has also been used to calculate the formation and transition energy of an interstitial/substitutional Mn defect and a self-interstitial in silicon [13, 14].

In this work we show that defect levels with similar orbital characters will cause the self-energy potential to be approximately zero, negating the DFT-$\frac{1}{2}$ correction. This is first illustrated on the tetrahedral hydrogen interstitial in diamond, an example chosen such that the problem is maximal. The goal of this work is to show why the self-energy potential is zero in these cases and to introduce a new DFT-$\frac{1}{2}$ technique to solve this problem: *The decoupled DFT-$\frac{1}{2}$ method for defects*. The hydrogen interstitial is first revisited using the new method. It is then shown that the band structure of the defect system can be reconstructed and it is compared to the band structure obtained from calculation with the HSE06 functional, a widely used functional in high quality defect calculations [15–18]. In addition to the hydrogen interstitial the negatively charged vacancy in diamond will also be studied with the decoupled DFT-$\frac{1}{2}$ method. The negatively charged vacancy was chosen because vacancies or vacancy related defects are a class of defects that likely require the decoupled DFT-$\frac{1}{2}$ method to calculate defect gaps. The reason for this is that the creation of a vacancy leaves behind dangling bonds from atoms of the host material, usually in some sort of symmetric configuration. This makes it likely that defect levels with similar orbital character appear.

A. The DFT-$\frac{1}{2}$ method

We now give a brief overview of the DFT-$\frac{1}{2}$ method, where we focus on those parts that are important to introduce the decoupled DFT-$\frac{1}{2}$ method. More details can be found in the original papers by Ferreira et al. [11, 19] or in the recent review of Mao et al. [7]. The DFT-$\frac{1}{2}$ method starts with Janak's theorem and the assumption that the Kohn Sham (KS) eigenvalue ε_α of orbital α is linearly dependent on the occupation of the orbital f_α, the validity of this assumption was verified in Ref. [20]. With this, one can derive that the band gap of a semiconductor, which is the difference between the ionization energy I and the electron affinity A, is equal to the difference between the eigenvalue of the half occupied conduction band minimum (CBM) and valence band maximum (VBM) or

$$\begin{aligned}\text{Band gap} &= I - A && (1)\\ &= \left(E_{tot}^{N-1} - E_{tot}^{N}\right) - \left(E_{tot}^{N} - E_{tot}^{N+1}\right) && (2)\\ &= \varepsilon_c(f_c = 1/2) - \varepsilon_v(f_v = -1/2) && (3)\end{aligned}$$

where the sub-indices c and v denote the CBM and VBM, respectively, and $f_\alpha = 0$ indicates that orbital α has the same occupation as in the ground state and $f_\alpha = \pm 1/2$ means that half an electron was added or subtracted from the ground state occupation. From Janak's theorem a new quantity can be derived: the *self-energy* S_α which is defined by

$$\frac{\partial \varepsilon_\alpha}{\partial f_\alpha} = 2S_\alpha. \tag{4}$$

By integrating Eq. (4), Eq. (3) can be rewritten in terms of the KS-gap and the self-energy:

$$\text{Band gap} = \text{KS-gap} + S_c + S_v \tag{5}$$

Equation (3) gives the impression that one could calculate the band gap with DFT by placing half an electron from the valence band in the conduction band. However, since the KS-eigenstates are Bloch states, which are delocalized, the self-energy of these states will be zero. Thus, according to (5) this approach gives no correction to the KS-gap or in other words, Bloch states do not accurately describe the localized holes [19]. Instead of changing the occupation within a calculation, a potential V_s is added to the pseudopotential as if the occupation was changed or as if the self-energy was added. This potential V_s is called the *self-energy potential* and the self-energy can be seen as a quantum mechanical average over this potential. The self-energy potential can be calculated as follows:

$$V_s = V_{KS}(f_\alpha = 0, \mathbf{r}) - V_{KS}\left(f_\alpha = -1/2, \mathbf{r}\right) \tag{6}$$

with V_{KS} the KS-potentials where the dependency on the electron density is not written explicitly and only the occupation of orbital α is considered as all other occupations remain the same. The KS-potentials of Eq. (6) are usually calculated for a single isolated atom using an all electron code. The Coulomb-like tail of the self-energy and the periodic boundary condition imposed in a DFT calculation will lead to a divergence. This divergence can be removed by defining a new self-energy potential $\tilde{V}_s(\mathbf{r}) = \Theta(\mathbf{r})V_s(\mathbf{r})$, where $\Theta(\mathbf{r})$ is a trimming function defined as

$$\Theta(\mathbf{r}) = \begin{cases} \left(1 - \left(\frac{r}{r_c}\right)^n\right)^3 & r \leq r_c \\ 0 & r > r_c \end{cases} \tag{7}$$

The trimming function introduces two new parameters n and r_c to the self-energy potential. The former is usually set to 8 as this gives a good balance between the cutoff sharpness and the potential smoothness [7]. The parameter r_c is called the cutoff radius and should be determined by extremizing the band gap [21]. This means that in order to calculate the band gap using DFT-$\frac{1}{2}$ one should sweep over multiple DFT calculations with self-energy potentials at different cutoff radii.

II. METHOD

A. The conventional DFT-$\frac{1}{2}$ method for defects

Since the DFT-$\frac{1}{2}$ method only uses DFT calculations the method has DFT computational scaling. This makes it an attractive method for defect calculations where a large supercell and a correct band gap are required. In Ref. [12] Lucatto et al. introduced the DFT-$\frac{1}{2}$ method for defect excitations. In this work, the DFT-$\frac{1}{2}$ method is used to calculate the gap between an occupied and unoccupied defect level in the band gap. This result can then be used to calculate the vertical transition energy between the ground and excited state of a defect, i.e. the absorption E_{abs} and emission energy E_{em}, as is also illustrated in Fig. 1. By calculating the Stokes or anti-Stokes shift, which must be done by plain DFT as this is a difference between total energies of two different structures, the zero phonon line (ZPL) can be obtained. The ZPL is an important and identifying property for color centers. Lucatto et al. [12] demonstrate this procedure for the NV^- center in diamond and find a ZPL of 1.84 eV close to the experimentally observed 1.95 eV [22].

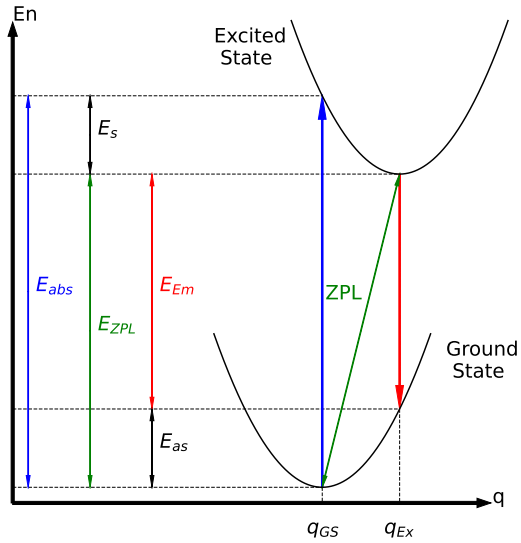

FIG. 1: A schematic overview of the energies involved in excitation and deexcitation of a defect. On the abscissas we have the configuration space q with the ground state configuration of the defect in its ground and first excited state denoted as q_{GS} and q_{Ex}, respectively. The blue and red lines are the absorption E_{abs} and emission energy E_{em}, respectively, and can be calculated using the DFT-$\frac{1}{2}$ method.

The DFT-$\frac{1}{2}$ method proposed by Lucatto et al., which we will call the conventional DFT-$\frac{1}{2}$ method for defects from now on, starts by dividing the atoms in the supercell into two groups: the defect atoms which are responsible for the defect levels and the bulk atoms which have a

negligible contribution to the defect levels and which are responsible for the valence and conduction bands. In order to use the DFT-$\frac{1}{2}$ corrected band gap in a defect calculation, all bulk atoms should use the pseudopotential determined by the DFT-$\frac{1}{2}$ method on the pristine host material of the defect. The group of defect atoms still has an unaltered pseudopotential allowing for an additional DFT-$\frac{1}{2}$ correction for the defect levels to be applied by adding a self-energy potential to these atoms. The general idea of the conventional method is the same, i.e. half an electron should be moved from the occupied to the unoccupied defect level in order to add self-energy. Instead of removing half the electron from the orbital with the largest contribution to the occupied defect level, a smaller fraction is removed from every orbital of every atom contributing to the defect level. The fraction removed from orbital ϕ of atom X is called

$$\xi_{X_\phi} = \frac{1}{2}\,\mathrm{char}_{X_\phi}[\psi_\alpha(\Gamma)] \tag{8}$$

with $\mathrm{char}_{X_\phi}[\psi_\alpha(k)]$ the projection of KS-state ψ_α of orbital α on to the atomic orbital ϕ of atoms X. Since exactly half an electron should be removed, ξ_{X_ϕ} should be normalized such that

$$\sum_{X_\phi} \xi_{X_\phi} = \frac{1}{2} \tag{9}$$

For the unoccupied level the fraction of electron which will be added to each orbital of each defect atom ζ_{X_ϕ} is calculated in a similar fashion. The self-energy potential for each orbital of each atom is then calculated as

$$V_S^{X_\phi} = V_X^{KS}(f_0 - \zeta_{X_\phi}) - V_X^{KS}(f_0 - \xi_{X_\phi}) \tag{10}$$

where only the occupation of the orbital ϕ is written as an input for V_X^{KS} as all other inputs are the same and with f_0 the ground state occupation of that atom. As is the case with bulk DFT-$\frac{1}{2}$, the self-energy potentials need to be multiplied by a trimming function Θ_{X_ϕ}. In the most general case, each orbital of each atom has its own trimming function and cutoff radius. However, we will assume that each orbital of the same atom has the same trimming function. Thus, the total self-energy potential of each atom is given by

$$V_S^X = \Theta_X \sum_\phi V_S^{X_\phi} \tag{11}$$

where Θ_X is given by Eq. (7). To find the cutoff of each atom, the gap between the defect levels needs to be extremized consecutively.

The problem that can appear in the conventional DFT-$\frac{1}{2}$ method is best illustrated on the tetrahedral hydrogen interstitial in diamond.

B. Computational details

The DFT calculations were performed in the local density approximation (LDA) exchange correlation potential and PAW [23] as implemented by the Vienna *ab initio* simulation package (VASP) [24–26] taking (collinear) spin polarization in to account. We chose LDA because Janak's theorem is exact for the LDA exchange correlation functional [19]. We calculate the lattice parameter of the conventional unit cell of diamond using a Birch-Murnagahan fit [27] with a cutoff energy of 520 eV and $8\times8\times8$ k-point grid in the Monkhorst-Pack scheme [28]. This gives us a distance between carbon atoms and a lattice parameter of 1.53 Å and 3.54 Å, respectively, which is in good agreement with the experimental values [29]. The defect supercells were created from a $4\times4\times4$ conventional diamond supercell with 512 carbon atoms. The integration over the Brillouin zone was done using only the Γ-point with an energy cutoff of 520 eV. Since the DFT-$\frac{1}{2}$ method does not produce a correct total energy, all relaxations were done using LDA. The relaxations were stopped when all forces are below 0.001 eV/Å.

The KS-potentials used for calculating the self-energy potential in Eq. (6) in the DFT-$\frac{1}{2}$ calculation were generated using a modified version of the ATOM code [11, 30]. The DFT-$\frac{1}{2}$ band gap for diamond was calculated by stripping 1/4th of both the s and p orbital as suggested by Ferreira et al. [11]. This results in a band gap of 5.73 eV and a cutoff parameter of 2.3 a_0, which is in line with the results obtained in Ref. [31, 32].

C. The hydrogen interstitial in diamond with the conventional method

In this section the electronic structure of the interstitial hydrogen defect in diamond is calculated using the conventional DFT-$\frac{1}{2}$ method. The nature of this defect makes it likely to have at least two defect levels with the same character, namely the $H_{s,\uparrow}$ and $H_{s,\downarrow}$ orbitals localized around the hydrogen atom. It is then demonstrated that the conventional method does not improve the DFT defect gap.

Interstitial hydrogen in diamond can either be found in a negative, positive or neutral charged state. We will focus on the neutral charged state, because in this charge state the defect level has one occupied and one empty defect level which is required by the conventional DFT-$\frac{1}{2}$ method. In the neutral state, interstitial hydrogen in diamond has 3 stable configurations. Going from lowest to highest energy these states are named BC (bond center), T (tetrahedral) and H (hexagonal) [33, 34]. Because the BC hydrogen defect was not suited for the DFT-$\frac{1}{2}$ method [35], the lowest meta-stable hydrogen interstitial, tetrahedral hydrogen, in diamond was studied. In this structure the hydrogen is located in one of the cavities of the diamond supercell. If this cavity and the hydrogen atom are placed along the [111] direction, then the hydrogen atom will be closer to one of the carbon atoms along this direction. If instead the hydrogen atom has the same distance to both carbon atoms along the [111] direction, then the defect is in the hexagonal state [33]. After re-

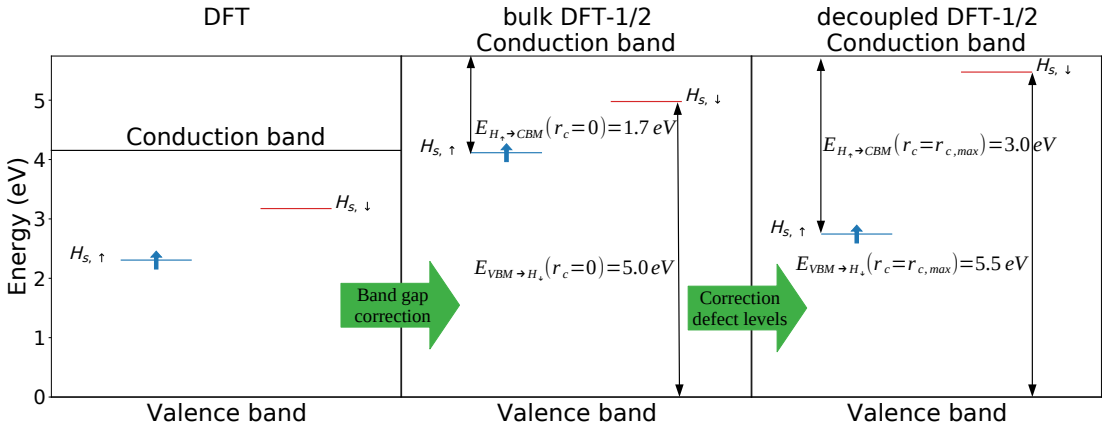

FIG. 2: The band structure at the Γ-point for the tetrahedral interstitial hydrogen in diamond, using LDA with an unmodified pseudopotential (left), LDA with the DFT-$\frac{1}{2}$ corrected pseudopotentials for bulk carbon atoms (middle) and the DFT-$\frac{1}{2}$ correction for bulk and defect atoms (right). The correction for the defect levels was obtained from the decoupled DFT-$\frac{1}{2}$ method. Because the decoupled method uses a different calculation for each defect level, the defect levels of the bulk DFT-$\frac{1}{2}$ band structure were shifted to their correct position with respect to the VBM and CBM as described in the text.

laxation the energy of the hydrogen interstitial is 1.09 eV higher than the energy of the BC configuration. This energy difference between the ground and meta-stable state is similar to that found in [33, 36] and deviates somewhat from [34], although this is likely because they use HSE06 instead of LDA.

In Fig. 2 the Γ-point band structure of the tetrahedral hydrogen defect is shown. In the case of tetrahedral hydrogen the only significant contribution to the defect levels of Fig. 2 comes from the s-orbital of the hydrogen atom, meaning that hydrogen is our only defect atom and $\xi_{H_s} = \zeta_{H_s} = 0.50$. With this the self-energy was determined with formula (10) and the gap between the defect levels was determined using the conventional DFT-$\frac{1}{2}$ method. The gap between the defect levels seems to be unaffected by the cutoff parameter, as can be seen in Fig. 3, the only difference is the run-to-run variance between the DFT runs. The reason for this failure of the conventional method will be explored in the next section.

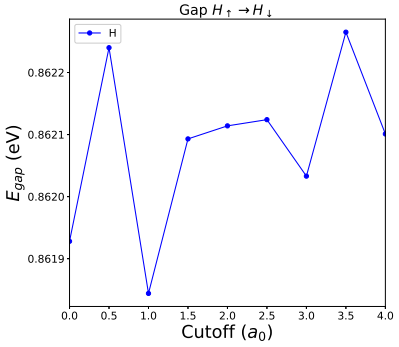

FIG. 3: The defect gap maximization for the interstitial tetrahedral hydrogen defect in diamond using the conventional DFT-$\frac{1}{2}$ method.

D. The decoupled DFT-$\frac{1}{2}$ method

The self-energy potential used in the DFT-$\frac{1}{2}$ method is spherically symmetric, meaning that only the s, p or d orbital character is looked at and the method does not differentiate between p_x, p_y and p_z for example. For some defects this results in a situation where $\xi_{X_\phi} \approx \zeta_{X_\phi}$, as was the case with the tetrahedral hydrogen interstitial. In these cases it follows that

$$V_X^{KS}(f_0 - \zeta_{X_\phi}) \approx V_X^{KS}(f_0 - \xi_{X_\phi}) \qquad (12)$$

By using Eq. (12) in (10) we see that both contributions to the self-energy potential for X_ϕ cancel:

$$V_S^{X_\phi} = V_X^{KS}(f_0 - \zeta_{X_\phi}) - V_X^{KS}(f_0 - \xi_{X_\phi}) \approx 0 \qquad (13)$$

which means that almost no self-energy is added when $\xi_{X_\phi} \approx \zeta_{X_\phi}$. This can be problematic when ξ_{X_ϕ} and ζ_{X_ϕ} are relatively large and thus a large part of the DFT-$\frac{1}{2}$ correction should come from the self-energy potential of orbital ϕ of atom X. Because the self-energy potential is approximately zero, almost no correction is added to the defect gap. To put it simply the conventional method does not work for these cases because it removes and then adds the same electron fraction to the defect atoms.

Generally, this cancellation of the self-energy po-

tential is the result of the approximation that the self-energy potential is spherical symmetric and thus this effect is unintended. To remove this cancellation of the self-energy the effect of $V_X^{KS}(f_0 - \xi_{X_\phi})$ should be decoupled from that of $V_X^{KS}(f_0 - \zeta_{X_\phi})$. This can be achieved by doing two separate calculations where either $V_X^{KS}(f_0 - \xi_{X_\phi})$ or $V_X^{KS}(f_0 - \zeta_{X_\phi})$ is added to the pseudopotential. Because $\xi_{X_\phi} \approx \zeta_{X_\phi}$ both the occupied and unoccupied level will move up or down together and the gap cannot be directly extremized as function of the cutoff. Instead, the defect gap should be extremized indirectly by extremizing $E_{VBM\to unocc}$: the gap between the VBM and the unoccupied level and $E_{occ\to CBM}$: the gap between the occupied level and the CBM. In Ref. [13] the cutoff parameter of Mn is determined in a similar fashion, by following the defect level in the DOS as a function of r_c with respect to the CBM. Both energy gaps are depicted in Fig. 4, where the electron transfer of the conventional and decoupled method are illustrated. Because both the valence and conduction bands have already had a bulk DFT-$\frac{1}{2}$ correction and the effect of the defect atoms on these bands is negligible, both of these bands can be used as reference bands.

These two gaps and the previously calculated DFT-$\frac{1}{2}$ band gap can formally be written as

$$E_{VBM\to unocc} = \varepsilon^+_{unocc} - \varepsilon^-_{VBM} \tag{14}$$
$$E_{occ\to CBM} = \varepsilon^+_{CBM} - \varepsilon^-_{occ} \tag{15}$$
$$E_{bandgap} = \varepsilon^+_{CBM} - \varepsilon^-_{VBM} \tag{16}$$

where E is used for energy gaps, ε for the KS-eigenvalues and the $+$ and $-$ super-indices are used to denote that these are the eigenvalues with half an electron added or subtracted, respectively. By using (14), (15) and (16) the formula of the defect gap $E_{gap,def}$ can be written in terms of the other energy gaps as follows:

$$\begin{aligned} E_{gap,def} =& \varepsilon^+_{unocc} - \varepsilon^-_{occ} && (17)\\ =& \left(E_{VBM\to unocc} + \varepsilon^-_{VBM}\right) \\ & - \left(\varepsilon^+_{CBM} - E_{occ\to CBM}\right) && (18)\\ =& E_{VBM\to unocc} + E_{occ\to CBM} \\ & - \left(\varepsilon^+_{CBM} - \varepsilon^-_{VBM}\right) && (19)\\ =& E_{VBM\to unocc} + E_{occ\to CBM} - E_{bandgap} && (20)\end{aligned}$$

And thus with Eq. (20) the gap between defect levels $E_{gap,def}$ has successfully been rewritten in terms of the energy gaps $E_{VBM\to unocc}$ and $E_{occ\to CBM}$ and the defect gap can be calculated based on two decoupled calculations. Because there are now two gaps and two cutoff parameters r_c to be determined the computational expense has doubled but the DFT scaling remains.

Since the defect gap in Eq. (20) depends on three quantities which are directly related to the band gap, any error made on this band gap by the bulk DFT-$\frac{1}{2}$ method will be carried over multiple times to the defect gap. We now determine the error of the defect gap as a result of

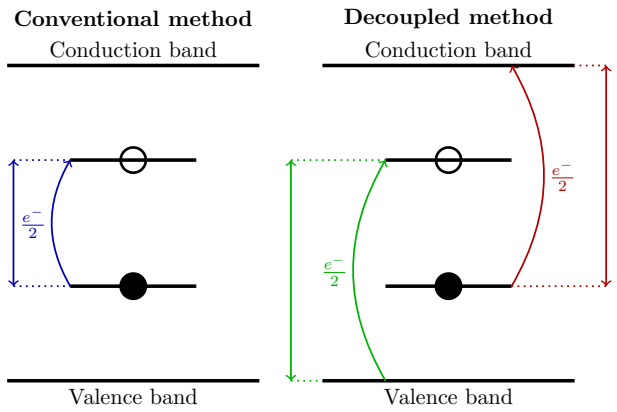

FIG. 4: A schematic of the KS-band structure with one occupied and unoccupied defect level depicted by the filled and unfilled circle, respectively. On the left the electron transfer of the conventional method is shown by the blue arrow. On the right the gaps $E_{VBM\to unocc}$ and $E_{occ\to CBM}$ as well as the corresponding electron transfers of the decoupled method are drawn by the green and red arrow, respectively.

the band gap error Δ, where we assume that the band gap error is the only source of error in the calculation, such that the energy eigenvalue of the defect level is bound by the energy eigenvalue of the defect level when the gap between this level and the VBM is exact and the energy eigenvalue of the defect level when the gap between it and the CBM is exact. The experimental band gap can be written as follows

$$E_{bandgap,exp} = E_{bandgap,DFT-1/2} + \Delta \tag{21}$$

where the sign of Δ can be either positive or negative. In what follows Δ is assumed to be positive but a similar expression can be obtained when Δ is negative. When $\Delta = 0$ the predicted DFT-$\frac{1}{2}$ gap and the actual gap are the same. In cases where $\Delta \neq 0$, the DFT-$\frac{1}{2}$ predicted gaps can have the following values:

$$E^{DFT-1/2}_{VBM\to unocc} \in [E_{VBM\to unocc}, E_{VBM\to unocc} + \Delta] \tag{22}$$
$$E^{DFT-1/2}_{occ\to CBM} \in [E_{occ\to CBM}, E_{occ\to CBM} + \Delta] \tag{23}$$

where the super-index $DFT - 1/2$ was added to denote the difference between the DFT-$\frac{1}{2}$ gap and the actual gap. Within this assumption the predicted values are either correct or at most a value $+\Delta$ off from the real values. This leaves 2 worst case scenarios where the error on $E_{gap,def}$ is maximal. The first being that both $E_{VBM\to unocc}$ and $E_{occ\to CBM}$ are predicted correctly leaving an error of $-\Delta$ coming from the band gap term in (20). The second happens when both $E_{VBM\to unocc}$ and $E_{occ\to CBM}$ have an error of Δ. One of these will be canceled by the error on the band gap while the other remains. These 2 worst case scenarios give the decoupled method an error margin of $\pm\Delta$ on $E_{gap,def}$ stemming from the band gap error.

III. RESULTS

A. The hydrogen interstitial in diamond with the decoupled method

Now that we have the tools to deal with cases where $\xi_{X_\phi} \approx \zeta_{X_\phi}$, we revisit the tetrahedral hydrogen interstitial in diamond. In Fig. 5 the cutoff parameters and sizes of the gaps for $E_{VBM\to H_{s,\downarrow}}$ and $E_{H_{s,\uparrow}\to CBM}$ are determined. With the maximum value of $E_{VBM\to H_{s,\downarrow}}$, $E_{H_{s,\uparrow}\to CBM}$ and Eq. (20) the gap between the $H_{s,\uparrow}$ and $H_{s,\downarrow}$ level was determined to be 2.7 eV, about 3 times larger than the gap of 0.9 eV when only the bulk DFT-$\frac{1}{2}$ correction is applied.

With the results of the decoupled DFT-$\frac{1}{2}$ method the band structure of the hydrogen interstitial was reconstructed starting from the bulk DFT-$\frac{1}{2}$ band structure. This bulk DFT-$\frac{1}{2}$ band structure is obtained from a DFT calculation where the bulk carbon atoms use the DFT-$\frac{1}{2}$ potential for pristine diamond and the defect hydrogen atom uses an unaltered pseudopotential. The bulk DFT-$\frac{1}{2}$ band structure obtained in this matter can be seen in the middle of Fig. 2. The occupied hydrogen level $H_\uparrow$ of this bulk DFT-$\frac{1}{2}$ calculation is then shifted down by $E_{gap}(r_c = r_{c,max}) - E_{gap}(r_c = 0)$ of the left curve of Fig. 5 such that the difference between the CBM and this level is exactly the maximum gap determined in this curve. The unoccupied level $H_\downarrow$ is shifted up in the same manner such that the gap between this level and the VBM is that of the curve on the right of Fig. 5.

The band structure has to be constructed in this way because the decoupled DFT-$\frac{1}{2}$ method uses two different pseudopotentials, one for each defect level. This means that there is no single calculation that can provide these eigenvalues and more importantly no wavefunction.

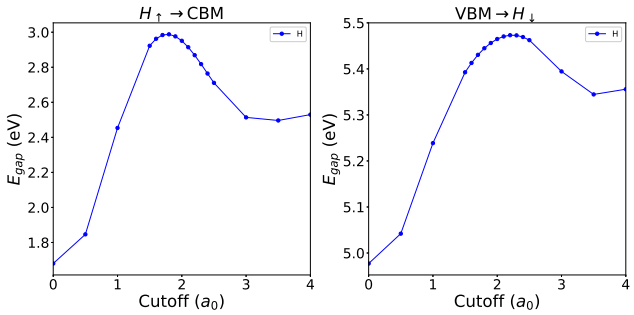

FIG. 5: The maximization of the gap between occupied spin up level and the CBM (left) and the VBM and the unoccupied spin down level (right) for $H_{int,T}$ used for the decoupled DFT-$\frac{1}{2}$ method.

To test whether the decoupled method produces a correct gap between the defect levels, the band structure for the hydrogen interstitial was calculated for a $2 \times 2 \times 2$ supercell using the decoupled method and a DFT calculation based on the HSE06 functional [9, 10]. Because this calculation is meant as a comparison between the electronic structure of the two different methods the same LDA relaxed supercell was used for both calculations. In Fig. 6 the HSE06 band structure is compared with both the bulk and bulk + the decoupled DFT-$\frac{1}{2}$ method. The band structure of the decoupled DFT-$\frac{1}{2}$ method on the right was again constructed by shifting the defect levels, the solid blue and red lines in Fig. 6, of the bulk DFT-$\frac{1}{2}$ band structure by $E_{gap}(r_c = r_{c,max}) - E_{gap}(r_c = 0)$ such that the gaps between these levels and the VBM and CBM are correct, according the decoupled method. Due to the small size of the supercell, the defect levels show some dispersion for both the DFT-$\frac{1}{2}$ and the HSE06 calculations. Since the HSE06 and the DFT-$\frac{1}{2}$ band gap for diamond differ by about 0.3 eV this is only a qualitative comparison. For both calculations we placed the zero-point of energy at the VBM. The HSE06 calculation finds a gap between defect levels of about 2.1 eV, which more closely matches the gap of 2.7 eV found by the decoupled method, as opposed to the gap found by the bulk DFT-$\frac{1}{2}$ method of 0.9 eV. The position of the defect levels with respect to the VBM and CBM of the decoupled method also matches the HSE06 result the closest. This leads us to conclude that if the conventional method fails, the decoupled method results in a qualitatively better result than applying no correction at all.

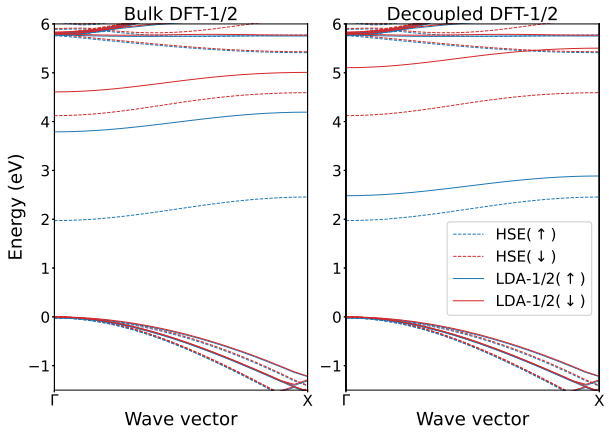

FIG. 6: The band structure of tetrahedral hydrogen for a $2 \times 2 \times 2$ diamond conventional supercell using bulk DFT-$\frac{1}{2}$ and HSE06 (left) and bulk+decoupled, DFT-$\frac{1}{2}$ and HSE06 (right). The zero point of energy was set at the VBM for all calculations.

Although the decoupled method was introduced as a solution for when the conventional method fails to determine the gap between two defect levels, it is demonstrated in Fig. 6 that the entire band structure can be corrected by calculating $E_{VBM\to unocc}$ and $E_{occ\to CBM}$ for all unoccupied and occupied defect levels, respectively.

B. The negatively charged vacancy in diamond

The negatively charged vacancy $V(-)$ consists of 5 electrons, 4 coming from the dangling carbon bonds and 1 from the extra negative charge. Experimentally it has been shown that in the ground state $V(-)$ has the $^4A_2$ many-body state with T_d symmetry [37]. Only the single-particle configuration $a_1(\uparrow,\downarrow)t_2(\uparrow,\uparrow,\uparrow)$ contributes to the $^4A_2$ may body state [15, 37, 38], which is stable against Jahn-Teller distortion [38]. The negatively charged vacancy is also responsible for the GR1 band and the corresponding ZPL at 3.15 eV [39]. This ZPL is associated with the transition from $^4A_2 \rightarrow {}^4T_1$, where the excited $^4T_1$ state can be written in terms of $a_1(\uparrow)t_2(\uparrow\downarrow,\uparrow,\uparrow)$ single particle states [15, 38]. Besides the negative charged state the vacancy can also be found in the neutral charge state with a ZPL of 1.67 eV [39], and theoretically it is also shown that the vacancy can be stable with a charge ranging from -2 to 2 [15].

In what follows we try to determine the ZPL of $V(-)$ with the DFT-$\frac{1}{2}$ method similar to what Lucatto et al. did for the NV^- [12]. We show that the conventional method of Lucatto et al. will not work for this defect and the decoupled method is required. However, the decoupled method also has difficulties with this defect. We provide an explanation for the problems that the decoupled method faces with this defect and offer a solution.

As a first step the electronic structure of $V(-)$ was calculated using DFT and bulk DFT-$\frac{1}{2}$, as depicted in Fig. 7. For the bulk DFT-$\frac{1}{2}$ calculation the four carbon atoms closest to the vacancy C_{vac} were treated as defect atoms. In the DFT-$\frac{1}{2}$ band structure both the a_1 and triple degenerate t_2 levels can be recognized, while the DFT band structure only has the triple degenerate t_2 levels. This makes the need for the band gap correction for this defect apparent, because without this correction the a_1 levels are hidden in the valence band.

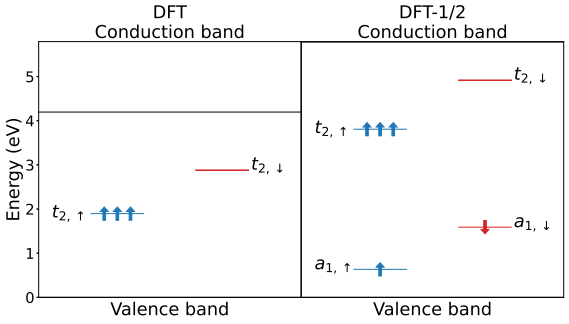

FIG. 7: The eigenvalues in the Γ-point for the negatively charged vacancy in diamond using DFT (left) and bulk DFT$-\frac{1}{2}$ (right), where the defect atoms are the 4 carbon atoms around the vacancy.

To calculate the ZPL one not only requires the ground state but also the first excited state (see Fig. 1), which is created by exciting an electron from $a_{1,\downarrow}$ to the $t_{2,\downarrow}$ state. The many-body excited state was approximated by placing one electron in one of the three triple degenerate $t_{2,\downarrow}$ levels using constrained DFT. Then the electron fractions ξ and ζ for the transition $a_{1,\downarrow} \leftrightarrow t_{2,\downarrow}$ were determined for both the ground and excited state, which can be found in table I. To obtain the defect gap correction only the decoupled method can be used since $\xi_{C,p} \approx \zeta_{C,p}$.

	ground state		excited state	
	ξ_{X_ϕ}	ζ_{X_ϕ}	ξ_{X_ϕ}	ζ_{X_ϕ}
$C_{vac,s}$	0.00	0.01	0.00	0.01
$C_{vac,p}$	0.12	0.11	0.12	0.12

TABLE I: The electron fraction for the transition $^4A_2 \leftrightarrow {}^4T_1$ of $V(-)$.

In what follows we focus on the calculation of the absorption energy with DFT-$\frac{1}{2}$ and the problems that come with it. The same story applies to the emission energy. In Fig. 8 the gaps $E_{VBM\rightarrow unocc}$ and $E_{occ\rightarrow CBM}$ are determined for the absorption energy. For the gap $E_{occ\rightarrow CBM}$ an unreasonably large cutoff of 5.6 a_0 was found, keeping in mind that the bond length between bulk carbon atoms is about 2.9 a_0 in our calculations. A cutoff this large encompasses the nearest neighbor of the C_{vac}, the entire spherical bulk self-energy of the nearest neighbor, the next nearest neighbors and the next next nearest neighbors. The absorption energy and the ZPL based on the maximum gaps of Fig. 8 are 4.0 eV and 3.9 eV, respectively. This deviates far from the previously mentioned experimental results. For the gap $E_{VBM\rightarrow unocc}$ a more reasonable cutoff of 3.2 a_0 was found.

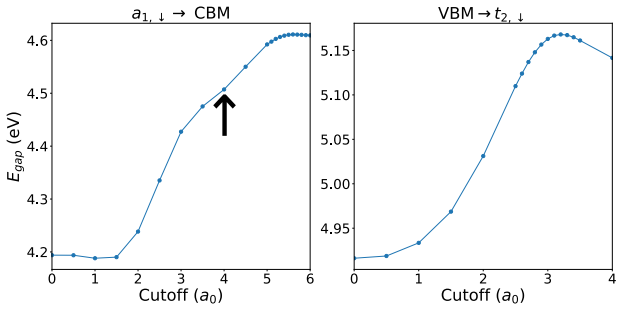

FIG. 8: The maximization of the gaps $E_{VBM\rightarrow unocc}$ and $E_{occ\rightarrow CBM}$ for the absorption energy of the $V(-)$ center. An arrow was added to the $E_{occ\rightarrow CBM}$ curve to highlight the kink in the curve.

We suspect that this unreasonably large cutoff is due to the other carbon atoms, that are located close to the vacancy and which also contribute to the occupied defect level. The curve of the gaps *occ to CBM* of Fig. 8 seems to have a kink at $r_c = 4.0$ a_0 due to two competing effects working on the defect gap, one with a maximum between 2.5 a_0 and 4.0 a_0 and one with a maximum around the global maximum of 5.6 a_0. The first maximum incorporates the nearest neighbors and most of the spatial re-

gion of its self-energy since the bulk cutoff is 2.3 a_0. The second maximum incorporates the next and next-next nearest neighbors of the C_{vac} atoms. It is as if adding self-energy to these neighboring atoms also increases the gap between the defect levels. It should be noted that the self-energy potential added to the neighbors of the C_{vac} atoms is not the correct self-energy potential for these atoms and these atoms already have a self-energy potential stemming from the bulk DFT-$\frac{1}{2}$ correction.

Upon closer inspection some other carbon atoms located near the vacancy have a nonzero contribution to the defect levels. If these atoms were to be included in the defect atom group they would have a ξ and ζ of the order of 0.01. The gaps $E_{VBM\to unocc}$ and $E_{occ\to CBM}$ where recalculated with the extra defect atoms in an attempt to prevent the large cutoff parameter by adding the correct self-energy potential to these atoms. However, this did not improve the results. Therefore, the neighboring atoms might not be the only reason for the large cutoff r_c or the incorrect self-energy was not the cause of the problem.

Since the cutoff parameter for $E_{occ\to CBM}$ cannot be found by maximizing this gap, the bulk parameter $r_c = 2.3\ a_0$ was used instead for both $E_{VBM\to unocc}$ and $E_{occ\to CBM}$. We argue that the cutoff parameter should be transferable and will not change much in different chemical environments, as is the case for the cutoff parameter in the bulk DFT-$\frac{1}{2}$ method [11]. Even in cases where the cutoff parameter can be calculated, taking a small deviation from this value by choosing the bulk cutoff parameter will not influence the gap greatly since close to the maximum the value of the gap is approximately constant. In Ref. [13] the cutoff parameter is determined for a Mn interstitial and substitutional defect in Si. This was done by following the energy of the defect level in the density of states with respect to the CBM as a function of r_c, similar to the *occ to CBM* in this work, resulting in a cutoff of 3.0 a_0 for Mn. This cutoff for Mn was also found in Ref. [40] where they use DFT-$\frac{1}{2}$ on GaMnAs. We further motivate this approach by determining the cutoff parameter using the conventional method. Since $\xi_{C,p} \approx \zeta_{C,p}$ in case of $V(-)$ but not $\xi_{C,p} = \zeta_{C,p}$, the effect on the gap of the DFT-$\frac{1}{2}$ approach is severely reduced but the method still produces the correct cutoff parameter of around 2.2 a_0 which is close to the bulk cutoff of 2.3 a_0.

The gaps $E_{VBM\to unocc}$ and $E_{occ\to CBM}$ were then calculated using the bulk cutoff parameter for both the ground and excited state such that the absorption and emission energy could be calculated using formula (20). With this the ZPL was calculated for the $V(-)$ center. The values of the optical transitions can be found in table II. In the best case we find a ZPL equal to the one found in [15] using HSE06. Our results seem to overestimate the optical transition energies, this is likely due to the approximated cutoff parameter and the error caused by the bulk DFT-$\frac{1}{2}$ correction which is about ± 0.23eV in the case of diamond.

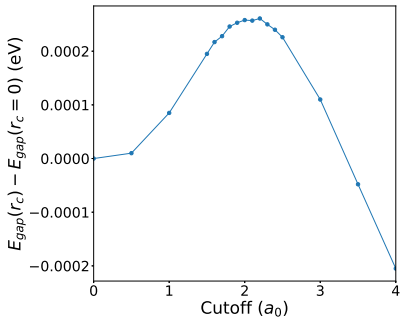

FIG. 9: The gap extremization of the absorption energy of $V(-)$ using the conventional DFT-$\frac{1}{2}$ method.

	optical transition	ZPL	shift (eV)
E_{abs}	3.64	3.47	0.16
E_{em}	3.08	3.30	0.22

TABLE II: The ZPL of $V(-)$ calculated based on the absorption and emission energy using the bulk cutoff parameter. The last column contains the Stokes and anti-Stokes shift depending on which is required to calculate the ZPL.

C. The general procedure of the decoupled DFT-$\frac{1}{2}$ method

We now give a brief overview of the decoupled DFT-$\frac{1}{2}$ method.

Preparation DFT-$\frac{1}{2}$

1. Use the DFT-$\frac{1}{2}$ method to correct the band gap of the host material.

2. Make the supercell with the defect and relax it with DFT.

3. Categorize all atoms in the supercell either as defect or bulk atoms.

4. Perform a scf calculation where the bulk atoms use the DFT-$\frac{1}{2}$ pseudopotential obtained in step 1, while the defect atoms use their unaltered pseudopotential.

5. Use the orbital contribution of each defect atom to the occupied and unoccupied level from the previous calculation to determine the electron fractions ξ_{X_ϕ} and ζ_{X_ϕ}.

6. If there are orbitals for which $\xi_{X_\phi} \approx \zeta_{X_\phi}$ use the decoupled DFT-$\frac{1}{2}$ method. Otherwise use the conventional method as described by Lucatto et al. [12].

The decoupled DFT-$\frac{1}{2}$ method

For each group of symmetrically equivalent atoms X (or orbitals X_ϕ) follow the steps below and use the DFT-$\frac{1}{2}$ pseudopotential obtained for these atoms while performing the calculations for the next group of atoms.

1. Make two sets of DFT-$\frac{1}{2}$ pseudopotentials, one where the electron fraction ξ_{X_ϕ} is added to X_ϕ for every value of the cutoff parameter r_c and one set where the ζ_{X_ϕ} is added.

2. Determine the optimal cutoff parameter for each set of pseudopotentials by running separate DFT calculations for each value of r_c in each set. The optimal cutoff parameter for ξ_{X_ϕ} is the one which extremizes the gap between the occupied defect level and the CBM. The optimal cutoff for ζ_{X_ϕ} should extremize the gap between the VBM and the unoccupied defect level.

3. Calculate the gap between the defect levels with the two extremes from the previous step and Eq. 20.

4. Repeat the process for the next defect atom (or orbital).

IV. CONCLUSIONS

To summarize and conclude, the conventional DFT-$\frac{1}{2}$ method for defects has been shown to fail in cases where the occupied and unoccupied defect level have a similar orbital character. This is due to the cancellation of the self-energy potential of the unoccupied level by the occupied level. When the electron fractions $\xi_{X,\phi}$ and $\zeta_{X,\phi}$ are large this can be problematic and one should use the decoupled DFT-$\frac{1}{2}$ method, which overcomes this problem by decoupling the effect of $\xi_{X,\phi}$ and $\zeta_{X,\phi}$. The decoupled method was tested on a tetrahedral hydrogen interstitial and the negatively charged vacancy in diamond. In the case of the hydrogen interstitial, it was shown that the decoupled DFT-$\frac{1}{2}$ method increases the gap between defect levels significantly which was qualitatively more in line with the HSE06 results. This comparison with HSE06 also showed that the decoupled DFT-$\frac{1}{2}$ method can be used to correct the entire band structure. For the negatively charged vacancy no proper cutoff parameters could be found because the localization of the defect atoms extended further than the nearest neighbors. Instead, the bulk cutoff parameter was used. The addition of this self-energy potential still leads to a ZPL similar to those calculated with HSE06. Although the decoupled method works in cases where the conventional DFT-$\frac{1}{2}$ method for defects fails it also has flaws. The decoupled method is sensitive to errors made by the bulk DFT-$\frac{1}{2}$ method, it sometimes does not find an appropriate cutoff as was the case with $V(-)$ and it doubles the computational cost.

ACKNOWLEDGEMENTS

This work was supported by the FWO (Research Foundation-Flanders), project G0D1721N. This work was performed in part using HPC resources from the VSC (Flemish Supercomputer Center) and the HPC infrastructure of the University of Antwerp (CalcUA), both funded by the FWO-Vlaanderen and the Flemish Government department EWI (Economie, Wetenschap & Innovatie).

[1] M. W. Doherty, N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup, and L. C. Hollenberg, Physics Reports **528**, 1 (2013).
[2] Ádám Gali, Nanophotonics **8**, 1907 (2019).
[3] G. Thiering and A. Gali, in *Diamond for Quantum Applications Part 1*, Semiconductors and Semimetals, Vol. 103, edited by C. E. Nebel, I. Aharonovich, N. Mizuochi, and M. Hatano (Elsevier, 2020) pp. 1–36.
[4] S. Pezzagna and J. Meijer, Applied Physics Reviews **8**, 011308 (2021).
[5] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).
[6] L. J. Sham and M. Schlüter, Phys. Rev. B **32**, 3883 (1985).
[7] G.-Q. Mao, Z.-Y. Yan, K.-H. Xue, Z. Ai, S. Yang, H. Cui, J.-H. Yuan, T.-L. Ren, and X. Miao, Journal of Physics: Condensed Matter **34**, 403001 (2022).
[8] J. Sun, A. Ruzsinszky, and J. P. Perdew, Phys. Rev. Lett. **115**, 036402 (2015).
[9] J. Heyd, G. E. Scuseria, and M. Ernzerhof, The Journal of Chemical Physics **118**, 8207 (2003).
[10] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, The Journal of Chemical Physics **125**, 224106 (2006).
[11] L. G. Ferreira, M. Marques, and L. K. Teles, Phys. Rev. B **78**, 125116 (2008).
[12] B. Lucatto, L. V. C. Assali, R. R. Pela, M. Marques, and L. K. Teles, Phys. Rev. B **96**, 075145 (2017).
[13] F. Matusalem, R. R. Pelá, M. Marques, and L. K. Teles, Phys. Rev. B **90**, 224102 (2014).
[14] F. Matusalem, M. Ribeiro, Jr., M. Marques, R. R. Pelá, L. G. Ferreira, and L. K. Teles, Phys. Rev. B **88**, 224102 (2013).
[15] P. Deák, B. Aradi, M. Kaviani, T. Frauenheim, and A. Gali, Phys. Rev. B **89**, 075203 (2014).
[16] A. Pershin, G. Barcza, Ö. Legeza, and A. Gali, npj Quantum Information **7**, 99 (2021).
[17] P. Deák, Physica B: Condensed Matter **535**, 35 (2018), 7th South African Conference on Photonic Materials

SACPM 2017.
[18] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, Rev. Mod. Phys. **86**, 253 (2014).
[19] L. G. Ferreira, M. Marques, and L. K. Teles, AIP Advances **1**, 032119 (2011).
[20] J. R. Leite and L. G. Ferreira, Phys. Rev. A **3**, 1224 (1971).
[21] In practise this usually means maximizing the band gap as DFT underestimates the band gap.
[22] G. Davies, M. F. Hamer, and W. C. Price, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **348**, 285 (1976).
[23] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[24] G. Kresse and J. Hafner, Phys. Rev. B **49**, 14251 (1994).
[25] G. Kresse and J. Furthmüller, Computational Materials Science **6**, 15 (1996).
[26] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
[27] F. Birch, Phys. Rev. **71**, 809 (1947).
[28] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
[29] M. Otfried, *Semiconductors - Basic Data*, 2nd ed. (Springer, 1996).
[30] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, Journal of Physics: Condensed Matter **14**, 2745 (2002).
[31] J. Doumont, F. Tran, and P. Blaha, Phys. Rev. B **99**, 115101 (2019).
[32] K.-H. Xue, J.-H. Yuan, L. R. Fonseca, and X.-S. Miao, Computational Materials Science **153**, 493 (2018).
[33] J. P. Goss, R. Jones, M. I. Heggie, C. P. Ewels, P. R. Briddon, and S. Öberg, Phys. Rev. B **65**, 115207 (2002).
[34] J. L. Lyons and C. G. Van de Walle, Journal of Physics: Condensed Matter **28**, 06LT01 (2016).
[35] As a first attempt the conventional DFT-$\frac{1}{2}$ method was tried on the BC configuration of the H interstitial i.e the lowest energy state. However, placing the hydrogen atom so close to the carbon atoms results in defect level with a non-negligible contribution from the p-orbitals of the carbon atoms that form the bond and the p-orbital of hydrogen. This means that these carbon atoms should be considered as defect atoms. The contribution of the H_p orbital is problematic for the DFT-$\frac{1}{2}$ method in general, because the method requires that a fraction of the empty p-orbital of the hydrogen atom in its ground state is removed. This problem will not be solved in this paper but could be the subject of a future study.
[36] A. Upadhyay, A. K. Singh, and A. Kumar, Computational Materials Science **89**, 257 (2014).
[37] J. Isoya, H. Kanda, Y. Uchida, S. C. Lawson, S. Yamasaki, H. Itoh, and Y. Morita, Phys. Rev. B **45**, 1436 (1992).
[38] S. J. Breuer and P. R. Briddon, Phys. Rev. B **51**, 6984 (1995).
[39] G. Davies, S. C. Lawson, A. T. Collins, A. Mainwood, and S. J. Sharp, Phys. Rev. B **46**, 13157 (1992).
[40] R. R. Pelá, M. Marques, L. G. Ferreira, J. Furthmüller, and L. K. Teles, Applied Physics Letters **100** (2012), 202408.