Abstract: In order to perform automated calculations of defect and dopant properties in semiconductors and insulators, we developed a software package, the Defect and Dopant ab-initio Simulation Package (DASP), which is composed of four modules for calculating: (i) elemental chemical potentials, (ii) defect (dopant) formation energies and charge-state transition levels, (iii) defect and carrier densities and (iv) carrier dynamics properties of high-density defects. DASP uses the materials genome database for quick determination of competing secondary phases when calculating the elemental chemical potential that stabilizes compound semiconductors. DASP calls the ab-initio software to perform the total energy, structural relaxation and electronic structure calculations of the defect supercells with different charge states, based on which the defect formation energies and charge-state transition levels are calculated. Then DASP can calculate the equilibrium densities of defects and electron and hole carriers as well as the Fermi level in semiconductors under different chemical potential conditions and growth/working temperature. For high-density defects, DASP can calculate the carrier dynamics properties such as the photoluminescence (PL) spectrum and carrier capture cross sections which can interpret the deep level transient spectroscopy (DLTS). Here we will show three application examples of DASP in studying the undoped GaN, C-doped GaN and quasi-one-dimensional SbSeI.

Key words: defect; dopant; first-principles calculations; carrier dynamics

Citation: M L Huang, Z N Zheng, Z X Dai, X J Guo, S S Wang, L L Jiang, J C Wei, and S Y Chen, DASP: Defect and Dopant ab-initio Simulation Package[J]. J. Semicond., 2022, 43(4), 042101. https://doi.org/10.1088/1674-4926/43/4/042101

1. Introduction

Intrinsic point defects are indispensable in crystals at a finite temperature and can have important influences on the fundamental properties of crystalline materials\cite{1,2}. Extrinsic elements are usually introduced into the crystalline lattice as dopants intentionally or as impurities unintentionally, which provides an extrinsic freedom for manipulating the properties of crystalline materials. The influences of defects and dopants (impurities) in semiconductors are more significant than those in metals or structural materials, because the applications of semiconductors in functional electronic and optoelectronic devices care more about the electrical and optical properties, which can be significantly changed even when a small number of defects and dopants exist in the lattice, e.g., defects or dopants with a density of $10^{15}$–$10^{18}$ cm$^{-3}$ (1 defect or dopant among $10^6$–$10^7$ atoms) can change the electrical and optical properties of semiconductors dramatically. For the simple elemental semiconductors such as Si and Ge, the number of intrinsic point defects is usually small\cite{3}, so extrinsic doping is required for producing electron or hole carriers and thus achieving n-type or p-type electrical conductivity. For binary, ternary, quaternary or even multinary compound semiconductors, the number of possible point defects can be large, so their defect/dopant physics can be more complicated and thus more flexible, e.g., their electrical conductivity may be tuned through controlling the formation of different point defects\cite{4}. Considering the important influences of defects (dopants) on semiconductor properties and the complexity of the defect/dopant physics in new, multinary or low-symmetry semiconductors, the studies on defects and dopants are fundamental to the development of semiconductor physics and applications.

Many experimental techniques have been developed for the characterization of defects and dopants in semiconductors, such as the photoluminescence (PL) spectrum, positron annihilation spectroscopy (PAS), electron paramagnetic resonance (EPR) and deep-level transient spectroscopy (DLTS)\cite{5}. These techniques usually just give the measured spectra, from which the defect and dopant properties, such as the density, energy levels and carrier capture cross sections can be obtained through fitting the spectra according to a certain physical model. However, what kind of defects and dopants determine the experimental spectra (including the fitted properties of defects and dopants) is usually speculated according to the analysis on the properties of the possible defects and dopants, which is semi-empirical and imposes a limit to the accurate and quantitative defect and dopant engineering. For instance, if the origin defects of the deep-level recombination centers in optoelectronic semiconductors are not clear, it is impossible to suppress their formation and increase the minority carrier lifetime by accurately controlling the growth conditions.

Since the 1990s, significant progresses have been made in the ab-initio (first-principles) calculation of defect and dopant properties based on the density functional theory (DFT) and the supercell model\cite{6–8}. The calculated formation energies, charge-state transition levels and carrier capture cross...
sections of defects and dopants have been widely used for revealing the microscopic origin of the experimental PL, PAS, EPR and DLTS spectra, and guiding the fabrication of high-quality materials and high-performance devices\cite{9−11}. However, there are still unsolved issues that may cause considerable errors in the calculated defect/dopant properties or even misunderstanding of the defect/dopant physics. (i) The approximations to exchange-correlation functionals cause the underestimation of the band gap and the incorrect location and occupation of the defect/dopant levels in the band gap\cite{12}, which can result in errors in the calculated formation energy and transition level. Such errors are serious for the local density approximation (LDA)\cite{13} and generalized gradient approximation (GGA)\cite{14}, so several correction schemes were proposed\cite{7, 15−18}. Recently, hybrid functional calculations were shown to predict more reasonable band gaps for many semiconductors, so the errors in the calculated defect properties are also corrected largely (not perfectly, because the hybrid functional may give incorrect localization of defect states and also depends on the empirical exact exchange ratio parameter)\cite{19−21}. (ii) The calculation of defect/dopant formation energies and densities requires the determination of the allowed chemical potential ranges for all the component and dopant elements. For ternary, quaternary and multinary compound semiconductors, a large number of secondary phases competing against the host material probably exist, which may all limit the stable range of elemental chemical potentials. Therefore, if some important secondary phases are not considered in the determination of elemental chemical potentials, the calculations of defect/dopant formation energies and densities and even the stability of the compound can be wrong. (iii) For low-symmetry and multinary compound semiconductors, the possible defect/dopant types can be many and various. For each type of defect/dopant, there may be multiple nonequivalent sites and structural configurations, which can change for different charge states. If some important defect types/sites/configurations/charge-states are not considered, the calculated results are incomplete and inaccurate. (iv) The finite supercell size corrections, such as the corrections for the electrostatic potential alignment and image charge interaction\cite{12, 22, 23}, are still controversial. Especially, different correction schemes can give rise to different results. (v) In most of the studies, only defect/dopant formation energies and transition levels were calculated, but the exact values of equilibrium defect densities, carrier densities and Fermi level are not calculated. Without the direct calculation of the defect densities, one may have misunderstandings on the influences of a certain defect, because all the defects (dopants) are correlated by the carrier densities and Fermi level, and the defect correlation may give rise to unexpected and anti-chemical-intuition changes in the defect densities\cite{24}.

To solve those problems in a standard and comprehensive way, we developed a software package, the Defect and Dopant ab-initio Simulation Package (DASP), which can perform automated calculations of defect and dopant properties based on the supercell model and ab-initio DFT calculations. Using the software, the defect formation energies, charge-state transition levels, carrier capture cross sections and even the PL spectrum can be calculated based on the atomic structure, total energy, electronic structure, phonon spectrum and electron–phonon coupling matrix calculated by the ab-initio DFT softwares using different approximations to exchange-correlation functionals. All the possible competing secondary compounds in the materials genome database are considered for the accurate calculation of the thermodynamic stability and elemental chemical potential ranges of compound semiconductors. Various defect types, atomic sites, structure configurations and charge states can be considered, and the corrections for the electrostatic potential alignment and image charge interaction can be included automatically. With all the defects and dopants considered, the equilibrium defect densities, carrier densities and Fermi level can be calculated for the samples grown under different chemical potential conditions and different temperature. For high-density defects and dopants, their carrier dynamics properties can also be calculated, despite the heavier computational cost. Therefore, DASP is developed to be a toolbox for the automated theoretical prediction of defect and dopant properties in semiconductors.

2. Software framework and modules

2.1. Software framework

DASP is composed of four modules: Thermodynamic Stability Calculation (TSC), Defect Energy Calculation (DEC), Defect Density Calculation (DDC), and Carrier Dynamics Calculation (CDC), as plotted in Fig. 1.

The detailed workflow of DASP is plotted schematically
in Fig. 2, in which the green part shows the DEC module, the blue part shows the TSC module, the purple part shows the DDC module and the red part shows the CDC module. The only necessary input is the crystal structure file of the semiconductor. The intermediate and final output files include:

(i) TSC: the chemical potential range of component elements that stabilizes the compound semiconductor (which is also a descriptor of the thermodynamic stability of the compound) and the allowed highest chemical potential of the dopant elements;

(ii) DEC: the formation energies of defects and dopants in different charge states, as functions of the elemental chemical potentials and Fermi level (electronic chemical potentials), from which the charge-state transition levels can be derived;

(iii) the equilibrium-state Fermi level, densities of electron and hole carriers, and densities of the defects and dopants in different charge states, as functions of the elemental chemical potentials at growth temperature and working (measuring) temperature;

(iv) the carrier capture cross sections, the radiative and non-radiative carrier recombination rates and lifetime, and the PL spectra.

Among the four modules, the most fundamental one is DEC, whose core function is for calculating the defect/dopant
formation energy using the supercell model and ab-initio DFT calculations. In the supercell model, the formation energy of a point defect \( \alpha \) in the charge state \( q \) can be calculated as\[^{25, 26}\]

\[
\Delta E_f(\alpha, q) = E(\alpha, q) - E(\text{bulk}) - \sum \mu_i(\mu_i + E_f) + q(E_F + \varepsilon_{\text{VBM}}) + E_{\text{corr}},
\]

where \( E(\alpha, q) \) and \( E(\text{bulk}) \) are the total energies of the supercells with and without a defect (dopant), \( \mu_i \) is the elemental chemical potential referenced to the total energy \( E_f \) of the pure solid/gas elementary phase, \( E_F \) is the Fermi level referenced to the eigenvalue of the valence band maximum (VBM) level of the bulk supercell. \( E_{\text{corr}} \) is the correction that accounts for the spurious interaction caused by finite supercell size and periodic boundary conditions.

In the following, we will briefly introduce the four modules:

**TSC module:** As shown in Eq. (1), the defect (dopant) formation energy and thus the defect density are functions of the elemental chemical potentials, so the calculation of defect properties requires the elemental chemical potentials as inputs. The TSC module is for calculating the chemical potential ranges through considering the influences of all the competing secondary compounds that can limit the thermodynamic phase stability of the host compound semiconductor. In order to make the consideration of secondary compounds as complete as possible, the material genome database (Materi- als Project\[^{27}\]) is used to search for the competing secondary compounds. Through combining the formation energy information of all the secondary compounds with the calculated formation energy of the host compound (equivalent inputs are used for DFT calculations), the critical secondary compounds that limit the stable chemical potential region can be determined quickly.

**DEC module:** The DEC module can generate the maxim- umly-cubic supercell structures of the host semiconductor, based on which the defect and dopant supercell structures can be generated with all the possible defect/dopant types, atomic sites and structural configurations considered. Then ab-initio software, such as Vienna Ab-Initio Simulation Package (VASP)\[^{28}\] and Quantum ESPRESSO (QE)\[^{29}\], will be called to perform atomic structural relaxation, total energy and electronic structure calculations for the defect/dopant supercells. First the neutral charge state is calculated and then different charge states will be generated according to the defect/dopant levels in the band gap and calculated. Then the formation energies of defects and dopants in different charge states can be calculated using Eq. (1). The chemical poten-tials outputted by TSC are used as \( \mu_i \). The correction term \( E_{\text{corr}} \) that accounts for the spurious interaction caused by finite supercell size and periodic boundary conditions can also be calculated automatically based on the output of ab-initio calculations. The charge-state transition levels are then calculated to be the Fermi level \( E_F \) at which the formation energies of a defect in two charge states \( q \) and \( q' \) are equal, \( \Delta E_f(\alpha, q) = \Delta E_f(\alpha, q') \).

**DDC module:** Based on the results of TSC and DEC, the DDC module solves the charge-neutrality equation self-consistently to determine the equilibrium defect/dopant densities, Fermi level, and carrier densities, as shown by the purple part of Fig. 2. For a given chemical potential condition, the formation energies of defects and dopants in different charge states are treated as input for a two-step self-consistent calculation, which solves the charge-neutrality condition equations at high growth temperature and low working (measuring) temperature to determine the Fermi level of the semi-conductor samples. Then the equilibrium densities of defects and dopants and the densities of electron and hole carriers can be calculated under the given chemical potential condition. Through changing the chemical potentials, the defect/ dopant and carrier densities can be calculated as functions of the chemical potential conditions, which can be used for the defect/dopant engineering based on the growth condition control.

**CDC module:** For the high-density defects or dopants identified by DDC module, the CDC module can calculate their excited-state carrier dynamics properties. The transition energy levels from the DEC module and the Fermi level from the DDC module will be used to screen for the transitions between the defect levels and the VBM or conduction band minimum (CBM) level that may produce the PL peaks or cause the fast non-radiative recombination of carriers. Then the phonon spectrum of the defect/dopant supercell, electron-phonon coupling matrix and transition dipole moments between the defect/dopant states and VBM or CBM states will be calculated, based on which the radiative and non-radiative transition rates (carrier capture cross sections) and the lineshape of photoluminescence spectra induced by defects and dopants can be calculated. Such calculations may be performed based on the single-phonon-mode one-dimensional configuration coordinate diagram\[^{30}\] or the all-phonon-modes three-dimensional schemes\[^{31}\]. Combining the calculated transition rates and the equilibrium defect and carrier densities, the CDC module may also calculate the Shockley–Read–Hall recombination rates and the non-equilibrium carrier lifetime.

### 2.2. Thermodynamic stability calculation module

The formation of defects or the doping of extrinsic elements in semiconductor lattice involves the exchange of atoms between the semiconductor lattice and the external environment, so the formation energies of defects and dopants in Eq. (1) depend on the elemental chemical potentials, which describe the abundance of the elements (partial pressure for the gas elements) in the environment. Although the abundance of a certain element can be controlled by changing the environment, the changes of the elemental chemical potentials are actually not unlimited, because they should satisfy a series of thermodynamic conditions to make the pure-phase crystalline semiconductor stable. Here we take the quaternary compound semiconductor \( \text{Cu}_2ZnSnS_4 \) as an example to show how to determine its stable elemental chemical potential range.

Firstly, at the equilibrium state of the compound \( \text{Cu}_2ZnSnS_4 \), the weighted sum of the chemical potentials of its component elements should be equal to the formation energy of the compound,

\[
2\mu_{\text{Cu}} + \mu_{\text{Zn}} + \mu_{\text{Sn}} + 4\mu_S = \Delta H_f(\text{Cu}_2ZnSnS_4),
\]

where \( \Delta H_f(\text{Cu}_2ZnSnS_4) \) is the compound formation energy of \( \text{Cu}_2ZnSnS_4 \) relative to the \( \text{Cu}, \text{Zn}, \text{Sn} \) and 5 elemental phases.
That means the chemical equilibrium state is arrived for the reaction $2\text{Cu} + \text{Zn} + \text{Sn} + 4\text{S} \rightarrow \text{Cu}_2\text{ZnSnS}_4$, so the product compound $\text{Cu}_2\text{ZnSnS}_4$ is stable.

Secondly, to make the synthesized sample be pure-phase $\text{Cu}_2\text{ZnSnS}_4$, the formation or coexistence of the competing secondary phases, such as the binary and ternary compounds $\text{CuS}, \text{Cu}_2\text{S}, \text{ZnS}, \text{SnS}, \text{SnS}_2$ and $\text{Cu}_2\text{SnS}_3$, and the elemental phases $\text{Cu}, \text{Zn}, \text{Sn}, \text{S}$, should be avoided. Therefore, the weighted sum of the chemical potentials of their component elements should be lower than their corresponding formation energies (the formation energies of elemental phases are 0), as described by the following inequations,

$$\mu_{\text{Cu}} + \mu_{\text{S}} < \Delta H_f(\text{CuS}),$$

$$2\mu_{\text{Cu}} + \mu_{\text{S}} < \Delta H_f(\text{Cu}_2\text{S}),$$

$$\mu_{\text{Zn}} + \mu_{\text{S}} < \Delta H_f(\text{ZnS}),$$

$$\mu_{\text{Sn}} + \mu_{\text{S}} < \Delta H_f(\text{SnS}),$$

$$\mu_{\text{Sn}} + 2\mu_{\text{S}} < \Delta H_f(\text{SnS}_2),$$

$$2\mu_{\text{Cu}} + \mu_{\text{Sn}} + 3\mu_{\text{S}} < \Delta H_f(\text{Cu}_2\text{SnS}_3),$$

$$\mu_{\text{Cu}} < 0,$$

$$\mu_{\text{Zn}} < 0,$$

$$\mu_{\text{Sn}} < 0,$$

$$\mu_{\text{S}} < 0.$$  \hspace{1cm} (3)

The allowed chemical potential range of $\text{Cu}, \text{Zn}, \text{Sn}$ and $\text{S}$ that stabilizes the pure-phase $\text{Cu}_2\text{ZnSnS}_4$ is limited by these equations and inequations. If the formation energies of the host and all the competing compounds are known, the ranges can be calculated and plotted in the chemical potential space, as shown in Ref. [4]. One example about SbSeI is also shown in Section 3.2.

The four elements $\text{Cu}, \text{Zn}, \text{Sn}$ and $\text{S}$ can form many binary, ternary and quaternary compounds, which can all act as the competing secondary phases of $\text{Cu}_2\text{ZnSnS}_4$. In principle, for the accurate calculation of the chemical potential range, one should take account of all the possible competing compounds, whose number can be large, especially for quaternary or multinary compounds. In DASP, the TSC module provides an automated solution for the quick screening of the critical competing phases and thus the accurate calculation of the chemical potential range. With the chemical formula of the compound, TSC will visit the materials genome database, such as the MP database, to search for all the compounds that are composed of the component elements, and then download the total energy, formation energy and DFT calculation input files of these compounds. Meanwhile, TSC will also perform ab-initio DFT calculation for the formation energy of the host compound with the input consistent with the MP database. With the formation energies of the host compound and all the competing compounds, TSC can solve the thermodynamic constraint equations and inequations to determine the stable chemical potential region and the critical competing compounds that limit the stable region directly. For the host and the critical competing compounds, TSC can further calculate their formation energies with higher accuracy and different functionals, to ensure that the calculated range of the chemical potentials are accurate. Usually these critical competing compounds are also those determining the energy above the convex hull, which shows the thermodynamic stability of the compound with respect to the phase separation into other competing compounds, so TSC can also be used for high-throughput and accurate calculation of the energy above convex hull and thermodynamic stability of compounds.

2.3. Defect energy calculation module

2.3.1. Maximumly-cubic supercell generation

In real semiconductor lattices, the densities of defects and dopants are usually much lower than the densities of atoms, e.g., there is only one defect or dopant among $10^4$–$10^7$ atoms for defects or dopants with a density of $10^{15}$–$10^{18}$ cm$^{-3}$. Therefore, the distance between two defects or dopants is usually very large. However, in the periodic supercell model, the distances between the defect and the periodic image defects in the neighboring supercell are actually not very large since the supercell has only several hundred atoms. Therefore, the supercell model causes unphysical interaction between defects or dopants, which may induce errors in the calculated defect properties.

To reduce the errors caused by the finite supercell size and achieve better supercell size convergence, DASP always tends to generate the supercells that are nearly cubic, so that the defect–defect (dopant–dopant) distance is maximized and the unphysical interaction is minimized. As illustrated in Fig. 3, for the two supercells with the same volume (same number of atoms), the smallest defect–defect distance in the nearly-cubic supercell should be larger than that in the supercell derived from direct expansion of primitive cell. In the nearly-cubic supercell, the smallest defect–defect distance should be along the (100), (010) or (001) direction, and the smallest distances along the three directions are similar. However, in the supercell derived from direct expansion of primitive cell, the smallest defect–defect distance may be along the (110) or (111) directions, and the distance can be smaller than the distance along the (100), (010) or (001) direction if the supercell basis vectors are not orthogonal and one of their angles is larger than 120°. Besides the advantage in maximizing the defect–defect distance, the nearly-cubic supercell also leads to the orthogonality of the reciprocal lattices, which may accelerate the ab-initio calculation and achieve better convergence.

The generation of the nearly-cubic supercell in DASP is through the maximumly-cubic supercell generation function, which firstly transforms the primitive cell into the nearly-orthogonal supercell and then expands it to get the nearly-cubic supercell which maximizes the cubic degree of the supercell with a given number of atoms.

To search for the nearly-orthogonal supercell, we developed a traversal searching method through trying different conversion matrices that can transform the lattice vec-
DASP adds two types of structural perturbations: (i) displacements of defects and dopants in the primitive-cell region of the supercell. For low-energy vacancy and antisite defects, there may be several non-equivalent atomic sites for one element, and the same-type defects on non-equivalent sites may have different properties. When generating vacancy, interstitial and antisite defects, DASP will consider all non-equivalent sites. For interstitial defects, DASP will search for the largest void region in the structure and meanwhile consider the Coulomb repulsion to determine the possible interstitial sites of cations and anions. For low-energy interstitial defects, DASP will automatically generate 10–20 different configurations randomly in the primitive-cell region of the supercell (different interstitial sites are ensured to be not close to each other).

For low-energy vacancy and antisite defects, there may be other distorted or metastable structure configurations, such as the DX centers\(^{32}\), which originate from large structural distortion of the antisite donor defects but act as acceptors after distortion. These distorted or metastable defects have been shown to have important impact on the electrical properties of semiconductors\(^{33–35}\). They can be obtained by imposing a structural perturbation to the locally relaxed structure of the defect in the charge state \(q\), because the perturbation may overcome the structural transition barrier to arrive at a new distorted structure.

DASP adds two types of structural perturbations: (i) displacing the atoms within a sphere around the defect by less than 0.5 Å randomly, as shown in Fig. 4(b); (ii) randomly distrib-

\[ \beta = \frac{V_{\text{supercell}}}{V_{\text{cube}}} + 0.001N_{\text{atom}}. \]  

The maximumly-cubic supercell is generated by DASP through maximizing the quantity \(\beta\) in a given range of \(N_{\text{atom}}\).
Fig. 4. (Color online) For a vacancy or antisite defect, the initial configuration (a) can be generated directly from the host lattice, and then structural perturbations are added, including (b) random displacements and (c) random distribution of the atoms within the sphere around the defect.

putting all the atoms within the sphere, as shown in Fig. 4(c). The default value of the sphere radius is set to 3 Å, but will be automatically increased to ensure at least 4 atoms in the sphere. The total number of the structures generated with the perturbation will be in the range of 10–20, but some of them may become identical after structural relaxation. Such kind of structural perturbations is in fact for achieving a global structural relaxation for the defect.

2.3.3. Charge state selection of ionized defects

Ionized defects in different charge states are fundamental to the basic understanding of defect physics, but there are still several open questions during the calculation of the properties of charged defects. The first one is how to determine the range of defect charge \( q \) for an ionized defect. In the past, the octet rule was often selected as a criterion to judge the charge range of a defect, which simply adopts the nominal charge of the single element. For instance, the sulfur vacancy should take 0, 1+, 2+ charges according to the simple octet rule, however, in non-conventional semiconductors such as MoS\(_2\), neither 1+ nor 2+ state of the sulfur vacancy is stable while the 1− charge state exists\(^{[30]}\).

To solve this problem, the estimation procedure of the charge range implemented in DASP is based on the calculated eigenvalues of the neutral defects at \( \Gamma \) point. Therefore, in the first step of DEC module, DASP will generate the structures of point defects in their neutral state, and then call ab-initio software to perform atomic relaxations and static self-consistent calculations. Once all the calculations are done, the eigenvalues of bulk and all the neutral defects will be extracted. The hybrid functional (such as HSE) is recommended for the static self-consistent calculations to obtain the more reasonable band gap and more reasonable location of the defect levels. As shown in Fig. 5, for defect A, if an occupied and an unoccupied level are found within the band gap between the VBM and CBM levels of the bulk, these two levels will be taken as defect levels and the defect A in 1+ and 1− charges will be calculated subsequently. Following this scheme, when three occupied levels and three unoccupied levels are found in the bulk band gap, the defect C in \([3+, 2+, 1+, 1−, 2−, 3−]\) charge states will all be calculated.

2.3.4. Electrostatic potential alignment

In Eq. (1), the eigenvalue of bulk VBM is used as the reference of the Fermi level, i.e., \( E_F = 0 \) means the Fermi level is located at VBM. However, the eigenvalues in the calculation of bulk and defect supercells can be compared only when the electrostatic potentials of the two periodic supercells are aligned\(^{[12]}\). This can be accomplished either by aligning the electrostatic potential far from the defect (e.g., the potential written in the LOCPOT file of VASP), or by aligning the deep core level of the farthest atom from the defect. The alignment term can be written as

\[
\Delta V_{q/b} = V(a, q) \mid_{\text{far}} - V(\text{bulk}) \mid_{\text{far}}.
\]  (7)

The correction of potential alignment to the formation energy of a defect in the charge state \( q \) is \( q \Delta V_{q/b} \), and is automatically included in the term \( E_{\text{corr}} \).

2.3.5. Image charge correction

The spurious Coulomb interaction between charged defect and its periodic images, as well as charged defect and the neutralizing background charge should be corrected from the calculated formation energies of ionized charged defects, which are called image charge corrections. DASP can currently perform the corrections according to two schemes. The first one is the Lany–Zunger scheme\(^{[12]}\), which can be represented by

\[
\Delta E_{\text{LZ}} = [1 - c_{\text{sh}}(1 - 1/\varepsilon)] q^2 a^2 / 2L^2.
\]  (8)

where \( a \) is the lattice-dependent Madelung constant, \( \varepsilon \) is the static dielectric constant, and \( L \) is the linear supercell dimension. Since DASP will always tend to generate a nearly-cubic supercell, the default value of \( c_{\text{sh}} \), is set to \(-1/3\)\(^{[12, 22]}\). Based on these definitions, the total correction term \( E_{\text{corr}} \) that appears in Eq. (1) can be separated into

\[
E_{\text{corr}} = q \Delta V_{q/b} + \Delta E_{\text{LZ}}.
\]  (9)

DASP also provides an interface with the code sxdefectalign written by Freysoldt, which implements the Freysoldt–Neugebauer–Van de Walle (FNV) scheme\(^{[23]}\) for charged defect correction. Different from the original paper that uses the planar average of the electrostatic potential to obtain the potential offset, we here use the atomic site potential instead, as Kumagai and Oba proposed\(^{[37]}\). In this scheme, the correction \( E_{\text{corr}} \) can be written as

\[
E_{\text{corr}} = E_{\text{lat}} - q(\Delta V_{q/b} - V_{\text{model}}).
\]  (10)

where \( E_{\text{lat}} \) is the Madelung energy and \( V_{\text{model}} \) is the potential obtained from the chosen model (such as Gaussian) charge density. Note that in the FNV scheme, the potential alignment defined in Sec. 2.3.4 is already included.

The two correction schemes mentioned above may be not valid for the charged defects in low-dimensional layered semiconductors, such as those in monolayer MoS\(_2\) or WSe\(_2\). A series of new correction schemes have been successfully de-
developed in the past decade for the charged defect correction in layered semiconductors\(^{[38–43]}\), which will be implemented in the future version of DASP.

### 2.3.6. Defect wavefunction initialization

The formation of a defect or dopant causes the change of the electronic wavefunction and the redistribution of charge density in the semiconductor lattice, which produces new electronic states (defect states) in the band gap. Although the wavefunction and charge density near the defect site can be changed dramatically, the area far from the defect should be weakly influenced. Therefore, for different defects or dopants in different charge states, the wavefunction and charge density in the region far from the defect site should be similar to those of the defect-free supercell. Since the self-consistent calculations of the wavefunction and charge density will be repeated hundreds of times, it will save a large amount of computational cost if we take advantage of the slightly changed wavefunction and charge density in the region far from the defect site to reduce the number of self-consistent calculation steps.

In DASP, the initial charge density of the neutral defect/dopant supercell is generated based on the converged wavefunction and charge density of the bulk host supercell, i.e., the charge density in the region far from the defect site is the same as that of the bulk host, while the charge density in the sphere around the defect is generated by a superposition of atomic charge densities and the charge density in the interface region is smoothed and rescaled according to the number of valence electrons of the defect supercell. Then the self-consistent calculations are performed to obtain the converged charge density and wavefunction of the neutral defect/dopant supercell. For the charged defect/dopant, its initial charge density is generated based on the converged wavefunction and charge density of the neutral defect through changing the occupation number of the defect states. With the initial charge density, the ab-initio self-consistent calculations can reach convergence with much less steps. The saving of computational cost can be more significant when large supercells are used.

### 2.4. Defect density calculation module

Under thermodynamic equilibrium, the density of a defect \(\alpha\) in the charge state \(q\) is determined by its formation energy according to\(^{[25, 44]}\)

\[
n(\alpha, q) = N_{\text{sites}} g_q e^{-\Delta E_{\text{d}}/(k_B T)},
\]

where \(N_{\text{sites}}\) is the density of the possible defect sites, \(g_q\) is the charge-dependent degeneracy factor, \(\Delta E_{\text{d}}\) is the defect formation energy at the given Fermi level and elemental chemical potentials as described by Eq. (1). All the ionized defects in the charge state \(q \neq 0\) produce carriers. The positively charged donor defects with \(q > 0\) produce electron carriers, and their summed charge is \(\sum_{\alpha, q>0} [\cdot n(\alpha, q)]\); while the negatively charged acceptor defects with \(q < 0\) produce hole carriers, and their summed charge is \(\sum_{\alpha, q<0} [\cdot (-n(\alpha, q))].\) The final densities of electron and hole carriers are contributed by both the thermal excitation and the ionization of all these defects (dopants). The equilibrium-state Fermi level can be calculated through solving the charge neutrality equation,

\[
n_0 + \sum_{\alpha, q>0} [\cdot (-q) \cdot n(\alpha, q)] = p_0 + \sum_{\alpha, q<0} [\cdot q \cdot n(\alpha, q)], \tag{12}
\]

where \(\sum_{\alpha, q>0} [\cdot (-q) \cdot n(\alpha, q)]\) and \(\sum_{\alpha, q<0} [\cdot q \cdot n(\alpha, q)]\) are the summed charges of negatively charged defects and positively charged defects, weighted by the charge \(q\). \(n_0\) and \(p_0\) are free carrier densities, which can be defined as

\[
n_0 = \int_{E_{\text{CBM}}}^{\infty} g_C(E) f(E) \, dE, \tag{13}
\]

\[
p_0 = \int_{-\infty}^{E_{\text{VBM}}} g_V(E) (1 - f(E)) \, dE, \tag{14}
\]

where \(g_C(E)\) and \(g_V(E)\) are the density of states (DOS) for the conduction and valence bands, respectively, and \(f(E)\) is the Fermi-Dirac occupation function. \(g_C(E)\) and \(g_V(E)\) can be calculated from both the parabolic band approximation and the exact integration of the ab-initio calculated band structure. The default setting in DASP is from the parabolic band approximation that requires the calculation of electron and hole effective masses, but it can be changed to the exact DOS from ab-initio calculations.

The semiconductors are usually grown or synthesized at a high temperature and then go through a rapid annealing process to a lower working (measuring) temperature. Therefore, the defects are usually formed at the high temperature and then the densities of different charge states will redistribute during the rapid annealing. The DDC module is developed in accordance with such fabrication process\(^{[44–46]}\) as shown schematically in Fig. 2. Eq. (12) is firstly solved at a high growth temperature, and then the Fermi level and the densities of each defect in different charge state \(q\) can be ob-
where \( e \) is the elementary charge, \( n \) is the density of carriers between two states and the lineshape function, which is mainly determined by the transition dipole moment \( \langle \alpha | \vec{r} | \beta \rangle \) of the initial and final states, and \( \hbar \omega_m \) and \( \hbar \omega_n \) are the corresponding eigenvalues. \( E_{ZPL} \) is corresponding to the charge-state transition level calculated in DEC module. At \( T = 0 \) K, all the charge carriers fall into the lowest-energy vibrational mode of initial state \( (m = 0) \), since there is no thermal activation for excited-state carriers. As a result, one can simply omit the summation over \( m \) in Eq. (15) to obtain the result for \( T = 0 \) K.

In order to evaluate the overlap integral in Eq. (15), one-dimensional configuration coordinate coordinate is used with an example given in Fig. 6(b), which adopts a single effective phonon mode to represent all \( 3N \) phonons in the system. This phonon mode can be depicted by the distortion of the defects between two charge states; the structural difference \( \Delta Q \) in the configuration coordinate can be given by

\[
\Delta Q = \sqrt{\sum_a m_a (R_{a\beta} - R_{a\alpha})^2},
\]

where \( R_{a\beta} \) and \( R_{a\alpha} \) are the Cartesian coordinates of the atom \( \alpha \) in initial and final supercell structures, and \( m_a \) is its mass. Using configuration coordinate diagram, the effective phonon frequencies \( \omega_m \) and \( \omega_n \) can be obtained by fitting the potential energy surfaces, and the integral can also be numerically calculated.

Configuration coordinate diagram is also useful for further calculating the radiative and nonradiative carrier capture coefficient of such process, which requires the exact calculation of transition dipole moment and electron-phonon coupling matrix element. In the current version, DASP only supports the one-dimensional static coupling method for calculating nonradiative carrier capture\(^{[30]} \), while the original static coupling theory including all \( 3N \) phonon modes will be implemented in the future version\(^{[31, 51]} \).

### 3. Calculation examples

In the following, we will show three examples about the application of DASP in the calculation of the defect and dopant properties of semiconductors, including the intrinsic point defect properties of the benchmark system GaN, the intrinsic point defect properties of the low-symmetry quasi-one-dimensional \( SbSe_2 \), and the PL spectrum of C-doped GaN.

#### 3.1. Intrinsic defects of GaN

GaN is a well-studied wide-band-gap semiconductor whose intrinsic defects have been studied by many groups since 1990s, so it is a good benchmark system for testing our DASP software. In Fig. 7, we show the calculated formation energies of vacancies, antisites and interstitials in GaN as functions of Fermi level, which are well consistent with the published results calculated using the hybrid functional\(^{[52–54]} \).

The formation energies of these point defects in neutral states are relatively high in GaN, no matter under Ga-rich or N-rich condition. Among them, the donor defect, N vacancy \( V_N \), has the lowest formation energy, but it still cannot pro-
produce a high density of electron carriers or good n-type conductivity and shift the Fermi level to close to the CBM level, even under the Ga-rich condition. Therefore, the formation of intrinsic point defects should not produce good n-type conductivity in GaN. For $V_{\text{N}}$, the DASP calculations based on the spin-polarized and hybrid functional ab-initio calculations show a (+/3+) transition level lying 0.4 eV above the VBM, and a (0/+) level close to the CBM. The (+/3+) transition level was usually not found in the early calculations based on the LDA or GGA to the exchange correlation functional[58], but was found in the recent hybrid functional calculations[52–54].

### 3.2. Intrinsic defects of SbSeI

Antimony selenoiodide (SbSeI) has a quasi-one-dimensional structure, as shown in Fig. 8(a), which is similar to that of Sb$_2$Se$_3$[24]. Its band structure is shown in Fig. 8(b), and has an indirect band gap of 1.79 eV. Its VBM is composed of the hybridized states of Sb-5s, Se-4p and I-5p, while the CBM is mainly Sb-5p. Using the TSC module, two secondary compounds, Sb$_2$I and Sb$_2$Se$_3$, are found to determine the allowed range of the Sb, Se and I elemental chemical potentials, together with the elemental phases of Sb, Se and I. Figs. 8(c) and 8(d) present the 3D and projected-2D phase stability diagrams in the elemental chemical potential space based on the results of the TSC calculations. The ternary compound SbSeI is stable only in the orange region surrounded by the competing phases Sb$_2$Se$_3$, Sb$_2$I, Sb and Se. The four boundary points, A, B, C and D, are selected as the chemical potential conditions for the following DEC and DDC calculations.

Fig. 9 shows the calculated formation energies of defects in SbSeI as functions of Fermi level for the four chemical potential conditions, A, B, C and D. The densities of these defects and carriers are also calculated using the DDC module and the results for the high-density defects are plotted in Fig. 10(a) as functions of the elemental chemical potential. Meanwhile, the change of Fermi level and carrier densities with the elemental chemical potential conditions is also plotted in Fig. 10(a).

It is obvious in Fig. 9 that there are many intrinsic defects with formation energies lower than 2 eV, so the formation of defects should be energetically easy in the quasi-one-dimensional lattice of SbSeI. Among them, the antisite defects between two anions, $V_{\text{Se}}$ and $I_{\text{Se}}$, should be the dominant defects with the lowest energy and highest density. $V_{\text{Se}}$ is an acceptor with a (0/-) level at 0.34 eV above VBM, which is not deep. $I_{\text{Se}}$ is a donor with a (0/+) level at 0.17 eV below CBM, which is relatively shallow. Figs. 10(b) and 10(c) plot the squared wavefunction of the Se$_{\text{I}}$ and I$_{\text{Se}}$ defect states, which show that the defect states are actually localized, although their corresponding transition levels are shallow, indicating the supercell size is sufficient for describing these defects in SbSeI. Besides them, there are also several other defects with formation energies lower than 1 eV and deep transition levels in the band gap, e.g., Se$_{\text{Se}}$ and $V_{\text{N}}$, which are both amphoteric with deep (0/+) and (0/-) transition levels.

Although the two dominant defects, the Se$_{\text{I}}$ acceptor and the I$_{\text{Se}}$ donor, have quite high density ($10^{17}–10^{18}$ cm$^{-3}$) and relatively shallow transition levels (should produce high density of electron and hole carriers), the final equilibrium density of carriers is not high due to the donor–acceptor compensation. Therefore, the Fermi level of the SbSeI sample with high densities of defects is always located near the middle of the band gap, i.e., at least 0.4 eV far from the CBM or VBM level as shown in Fig. 11(a), and the electrical conductivity can only be weakly n-type or weakly p-type, regardless of the elemental chemical potential.

In contrast with the shallow Se$_{\text{I}}$ and I$_{\text{Se}}$ whose densities are always high, the densities of deep-level defects Se$_{\text{Se}}$ and $V_{\text{N}}$ change significantly as the elemental chemical potentials change. When the chemical potential points are close to the A-D line, the density of $V_{\text{N}}$ is very high; when the points are close to B-C, its density decreases to lower than $10^{15}$ cm$^{-3}$. For Se$_{\text{Se}}$, its density is very high when the chemical potential point is near the C point, but decreases quickly as the point moves away from the C point. These results suggest that the chemical potential should be controlled at the intermediate region of the B-C path, in order to suppress the formation of deep-level defects such as Se$_{\text{Se}}$ and $V_{\text{N}}$.

As we can see, Figs. 8–10 give a good example about the application of DASP for calculating the elemental chemical potential region that stabilizes the compound semiconductors, then the formation energies and transition levels of all intrinsic point defects, and finally the densities of defects and electron/hole carriers. For new semiconductors, the similar calculations should be performed to have a complete picture about their defect and dopant properties.

### 3.3. PL spectrum of C-doped GaN

The PL spectrum is a widely used optical characterization method of defects in semiconductors. Usually the origin of the PL peaks is attributed to different defects according to the energy differences between the defect level and band edge levels. However, such kind of attribution is questionable, especially when there are several defects whose energy levels are nearly degenerate. DASP can calculate the PL lineshape of different defects, which provides an extra criterion for the identification of the defect origin of PL peaks.

Here we show an example about the calculated PL lineshape of the $C_{\text{N}}$ dopant in GaN. Using DEC module, we find $C_{\text{N}}$ has a (0/+) transition level of 0.39 eV, and a (0/-) transition level of 1.07 eV above the VBM level, which agree with the calculations of Lyons et al.[57]. Therefore, the photoexcited electron carrier can be captured by the (0/-) level of $C_{\text{N}}$ through a radiative transition, from the initial state of neutral $C_{\text{N}}$ ($C_{\text{N}}^0$) with an electron on the CBM to the final state of neg-
atively charged $C\mathrm{N}^-$ with the electron on the (0/–) defect level. This transition is plotted in the configuration coordinate diagram in Fig. 11(b). The direct optical transition level is calculated to be 2.11 eV. The PL lineshape of the transition calculated using the CDC module is plotted in Fig. 11(c), which is in good agreement with experimental measurement and previous calculated result [50].

Further CDC calculation shows the radiative carrier capture coefficient $C_c$ is $1.7 \times 10^{-13}$ cm$^3$/s, slightly larger than the previous calculated result of $0.7 \times 10^{-13}$ cm$^3$/s [58]. The main reason is attributed to the larger momentum matrix element that we calculated, which is $|p_{\text{if}}|^2 = 0.06$ a.u. between the initial and final states.

4. Conclusions

We developed a software DASP composed of four mod-
ules, TSC, DEC, DDC and CDC, for the automated calculation of defect and dopant properties in the crystalline semiconductors (insulators). DASP just needs the input of the crystal structure file of the semiconductor, then it can visit the materials genome database and call the ab-initio DFT software such as VASP to calculate the defect and dopant properties, including, (i) the chemical potential range of component elements that stabilizes the compound semiconductor (a descriptor of the thermodynamic stability of the compound) and the highest allowed chemical potential of dopant elements; (ii) the formation energies of defects and dopants in different charge states, as functions of elemental chemical potentials and Fermi level, and their charge-state transition levels; (iii) the equilibrium densities of defects and dopants, Fermi level, densities of electron and hole carriers, as functions of elemental chemical potentials at growth and working temperature; and (iv) the carrier capture cross sections, radiative and non-radiative carrier recombination rate and lifetime, and the PL spectra. DASP is designed to be an automatic toolbox for the theoretical calculation studies on defects and dopants in semiconductors, and can be used not only for interpreting the electrical and optical characterization experiments of defects and dopants, but also for the quantitative defect and dopant engineering in functional semiconductor devices.

Acknowledgements

We thank Profs. Suhuai Wei, Xingao Gong, Aron Walsh, Linwang Wang and Yuning Wu, and Drs. Zhenkun Yuan, Jiqiang Li, Zenghua Cai and Tao Zhang for their long-term collaboration and very helpful discussion. The development of the commercial version of DASP is supported by the joint project between Hongzhiwei Technology (Shanghai) Co., Ltd. and Fudan University.

References

[1] Pantelides S T. The electronic structure of impurities and other point defects in semiconductors. Rev Mod Phys, 1978, 50, 797
[2] Park J S, Kim S, Xie Z J, et al. Point defect engineering in thin-film solar cells. Nat Rev Mater, 2018, 3, 194
[3] Ramprasad R, Zhu H, Rinke P, et al. New perspective on formation energies and energy levels of point defects in nonmetals. Phy Rev Lett, 2012, 108, 066404
[4] Chen S Y, Walsh A, Gong X G, et al. Classification of lattice defects in the kesterite Cu2ZnSnS4 and Cu2ZnSnSe4 earth-abundant solar cell absorbers. Adv Mater, 2013, 25, 1522
[5] Alkauskas A, McCluskey M D, van de Walle C G. Tutorial: Defects in semiconductors—Combining experiment and theory. J Appl Phys, 2016, 119, 181101
Giannozzi P, Baroni S, Bonini N, et al. QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials. J Phys: Condens Matter, 2009, 21, 395502

Alkauskas A, Yan Q M, van de Walle C G. First-principles theory of nonradiative carrier capture via multiphonon emission. Phys Rev B, 2014, 90, 075202

Shi L, Xu K, Wang L W. Reply to “Comment on ‘Comparative study of ab initio nonradiative recombination rate calculations under different formalisms’”. Phys Rev B, 2018, 97, 077302

van de Walle Chris G. DX-center formation in wurzite and zinc-blende Al,Ga$_{1-x}$N. Phys Rev B, 1998, 57, R2033

Park C H, Zhang S B, Wei S H. Origin of p-type doping difficulty in ZnO: The impurity perspective. Phys Rev B, 2002, 66, 073202

Li J B, Wei S H, Wang L W. Stability of the DX–center in GaAs quantum dots. Phys Rev Lett, 2005, 94, 185501

Li J L, Yang J X, Wu T, et al. Formation of DY center as n-type limiting defects in octahedral semiconductors: The case of Bi-doped hybrid halide perovskites. J Mater Chem C, 2019, 7, 4230

Komsa H P, Krasheninnikov A V. Native defects in bulk and monolayer MoS$_2$ from first principles. Phys Rev B, 2015, 91, 125304

Kumagai Y, Oba F. Electrostatics-based finite-size correction for first-principles point defect calculations. arXiv: 1402.1226, 2014

Wang D, Han D, Li X B, et al. Determination of formation and ionization energies of charged defects in two-dimensional materials. Phys Rev Lett, 2015, 114, 196801

Komsa H P, Pasquarello A. Finite-size supercell correction for charged defects at surfaces and interfaces. Phys Rev Lett, 2013, 110, 095505

Wu Y N, Zhang X G, Pantelides S T. Fundamental resolution of difficulties in the theory of charged point defects in semiconductors. Phys Rev Lett, 2017, 119, 105501

Xiao J, Yang K K, Guo D, et al. Realistic dimension-independent approach for charged-defect calculations in semiconductors. Phys Rev B, 2020, 101, 165306

Zhu G J, Yang J H, Gong X G. Self-consistently determining structures of charged defects and defect ionization energies in low-dimensional semiconductors. Phys Rev B, 2020, 102, 035202

Freyssoldt C, Neugebauer J. First-principles calculations for charged defects at surfaces, interfaces, and two-dimensional materials in the presence of electric fields. Phys Rev B, 2018, 97, 205425

Ma J, Wei S H, Gessert T A, et al. Carrier density and compensation in semiconductors with multiple dopants and multiple transition energy levels: Case of Cu impurities in CdTe. Phys Rev B, 2011, 83, 245207

Yang J H, Park J S, Kang J, et al. Tuning the Fermi level beyond the equilibrium doping limit through quenching: The case of CdTe. Phys Rev B, 2014, 90, 245202

Wei J C, Jiang L L, Huang M L, et al. Intrinsic defect limit to the growth of orthorhombic HFO$_2$ and (HF,Zn)O$_2$ with strong ferroelectricity: First-principles insights. Adv Funct Mater, 2021, 31, 2104913

Stoneham A M, Smoluchowski R. Theory of defects in solids: Electronic structure of defects in insulators and semiconductors. Phys Today, 1976, 29, 62

Huang M L, Wang S S, Wu Y N, et al. Defect physics of ternary semiconductor ZnGeP$_2$ with a high density of anion-cation antisties: A first-principles study. Phys Rev Appl, 2021, 15, 024035

Schanovsky F, Göss W, Grasser T. Multiphonon hole trapping from first principles. J Vac Sci Technol B, 2011, 29, 01A201

Alkauskas A, Lyons J L, Stein a Y, et al. First-principles calculations of luminescence spectrum line shapes for defects in semiconductors: The example of GaN and ZnO. Phys Rev Lett, 2012, 109, 267401

Cai L L, Wang S S, Huang M L, et al. First-principles identification of deep energy levels of sulfur impurities in silicon and their carrier capture cross sections. J Phys D, 2021, 54, 335103

M L Huang et al.: DASP: Defect and Dopant ab-initio Simulation Package
Menglin Huang obtained his Ph.D. from Department of Electronics, East China Normal University. Currently he is a postdoc in School of Microelectronics, Fudan University. His research focuses on the defect and dopant physics in semiconductor materials and devices. He is now working on the first-principles simulation of charge trapping by defects in different field-effect-transistor technologies.

Shiyou Chen is a professor in School of Microelectronics, Fudan University. He obtained his Ph.D. from Fudan University in 2009, and then worked in East China Normal University and Lawrence Berkeley National Laboratory during 2009–2020. His research focuses on the first-principles studies on defect (dopant) physics, multinary compound semiconductors and reliability physics of semiconductor devices.