Computational Materials Science 130 (2017) 1–9

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Editor's Choice

A computational framework for automation of point defect calculations

Anuj Goyal [a,b,*], Prashun Gorai [a,b], Haowei Peng [b], Stephan Lany [b], Vladan Stevanović [a,b]

[a] Colorado School of Mines, Golden, CO 80401, USA
[b] National Renewable Energy Laboratory, Golden, CO 80401, USA

ARTICLE INFO

Article history:
Received 14 November 2016
Received in revised form 28 December 2016
Accepted 30 December 2016

Keywords:
Point defects
High-throughput
Density-functional theory
Finite-size corrections
Materials genome initiative

ABSTRACT

A complete and rigorously validated open-source Python framework to automate point defect calculations using density functional theory has been developed. The framework provides an effective and efficient method for defect structure generation, and creation of simple yet customizable workflows to analyze defect calculations. The package provides the capability to compute widely-accepted correction schemes to overcome finite-size effects, including (1) potential alignment, (2) image-charge correction, and (3) band filling correction to shallow defects. Using Si, ZnO and In_2O_3 as test examples, we demonstrate the package capabilities and validate the methodology.

Published by Elsevier B.V.

1. Introduction

In semiconductor materials, point defects play a vital role in determining their properties and performance, particularly in microelectronics [1], optoelectronics [2], and thermoelectrics [3] related applications. The dominant point defects and their concentrations are determined from the defect formation energies, which can be predicted with reasonable accuracy [4] using first-principles methods such as density functional theory (DFT). Therefore, computational modeling of point defects is increasingly becoming an indispensable tool to understand and predict behavior of semiconductors [5–7]. Modern approaches to point defect calculations uses DFT and are typically based on the supercell approach [6,7]. With the goal of accelerating the design and discovery of materials by large-scale deployment of defect calculations, we have developed a computational framework (Fig. 1) to automate supercell-based point defect calculations with DFT. Our approach successfully addresses two main challenges of automating point defect calculations: (1) generation of defects structures including vacancies, substitutional defects and interstitials, and (2) application of the finite-size and band gap corrections.

In the context of structure generation, creating supercells with vacancies and substitutional defects is relatively straightforward. In contrast, identifying likely locations of interstitials is much more challenging because of the large number of interstitialcy sites, especially in complex, multinary systems. In addition, interstitials might adopt complex configurations, including the split or dumbbell where the interstitial is associated with a off-site lattice atom. To address these challenges, we have developed an efficient scheme based on Voronoi tessellation; [8] the scheme considers corners, edge and face centers of the Voronoi polyhedra as likely sites for interstitials. We demonstrate that, upon relaxing the structure, this scheme successfully discovers both the symmetric and general Wyckoff positions as well as the split interstitial configurations. Our implementation of this scheme is independent of pymatgen [9] where Voronoi tessellation is also employed. Here we will discuss the algorithm in detail and validate the Voronoi-driven approach to identify interstitial sites.

Within the supercell approach to calculate the defect formation energies, finite-size artifacts need to be removed using carefully designed correction schemes. We have implemented tools to calculate the following finite-size corrections: (1) potential alignment, (2) image-charge correction, and (3) band filling correction to address Moss-Burstein-type effects. We follow the widely used and tested approach of Lany and Zunger [10,11] out of the several others that addresses the same issues [12–16]. However, the automated framework is highly modular so that other correction schemes can be easily implemented including computation of defect formation energies using series of supercell sizes in order to extrapolate the values to the infinitely large supercell. In addition, the framework employs fitted elemental-phase reference energies (FERE) [17,18] to compute elemental chemical potentials, which have been shown to provide accurate predictions of thermodynamic phase stability.

* Corresponding author at: Colorado School of Mines, Golden, CO 80401, USA.
E-mail addresses: anuj.goyal@nrel.gov (A. Goyal), Vladan.Stevanovic@nrel.gov (V. Stevanović).

http://dx.doi.org/10.1016/j.commatsci.2016.12.040
0927-0256/Published by Elsevier B.V.

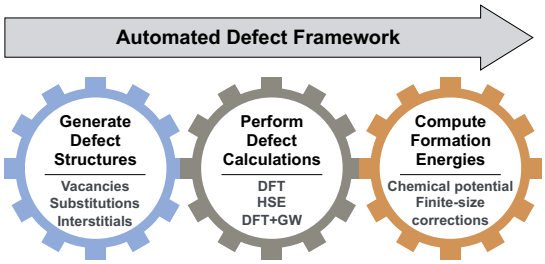

Fig. 1. Three key components of the computational framework to automate point defect calculations.

Beyond the finite-size effects, another source of inaccuracy arises from the well-known DFT band gap problem. Accurate band gaps are needed to correctly describe the formation energy of charged defects as a function of the electronic chemical potential i.e., Fermi energy. We employ state-of-the-art GW quasiparticle energy calculations [19] to compute band edge shifts (relative to the DFT-computed band edges). The band edge shifts are used to correct the defect formation energy in multiple charge states. The automated framework is also capable of performing defect calculations with DFT hybrid functionals [20,21]. However, supercell defects calculations with hybrid functionals have sources of uncertainty arising from the choice of input parameters (e.g. fraction of exchange) and have considerable computational overheads [22]. Therefore, we have implemented a DFT+GW approach for calculating defect formation energy that has been shown to be as accurate as calculations with hybrid functionals [22].

Finally, we illustrate and validate the automated computational framework by considering the set of three well-studied semiconductor materials, Si, ZnO, and In_2O_3 with a total of 17 unique interstitial and vacancy structures in multiple charged states. We show that our results on defect formation energies and charge defect transition levels in Si, ZnO and In_2O_3 agree well with the literature. The framework successfully identifies the known intrinsic interstitial and vacancy structures in Si, ZnO and In_2O_3. In addition, it discovers interstitial structures in In_2O_3, with formation energies ~0.5 eV above that of previously known interstitial structures.

2. Overview of the automated defect framework

Fig. 2(a) presents a workflow of the automated framework, including generation of defect structures, relaxation of defect supercells within DFT using the PyLada framework [23], and determination of finite-size and band gap corrections to compute the defect formation energies. In this section, we describe each component of the framework and provided technical details. The latest version of the package can be downloaded from GitHub repository at https://github.com/pylada/pylada-defects.

2.1. Generate defect structures

The workflow takes the fully-relaxed primitive cell as an input to create supercells. To create a vacancy or substitutional defect in supercell, the occupied Wyckoff positions (lattice-sites) for all atom types in the supercell are identified. Then the corresponding atom is removed or substituted with an impurity atom, to generate vacancy or substitutional defect. Finally, the first nearest-neighbor atoms to the vacancy or substitutional site are randomly displaced (~0.1 Å) to break the underlying site symmetry and thereby, ensuring the non-symmetric configurations of the defects are properly captured. The Voronoi tesselation [8,24], scheme is employed to identify likely interstitial sites. Voronoi region is the volume that encloses the points p closest to a given lattice site P_i than to any other lattice site P_j for $i,j \in I_n = \{1,\ldots,n\}$. Mathematically, it is defined as [8]

$$V(P_i) = \{p | d(p,P_i) \leqslant d(p,P_j)\} \quad \text{for } j \neq i,\ j \in I_n \tag{1}$$

where $V(P_i)$ is the Voronoi region associated with P_i, and $d(p,P_i)$ specifies the minimum distance between a general point p and P_i. To create an interstitial, Voronoi regions (Eq. (1)) are computed across each occupied Wyckoff positions, and symmetry inequivalent vertices, face, and edge centers of the Voronoi regions are chosen as the candidate sites for the interstitials. The number of candidate interstitial sites depends on the symmetry of the crystal structure. The lower the symmetry and the more complex the crystal structure, the larger the number of sites. For example, in In_2O_3 (space group *Ia*-3, 40 atoms in primitive cell), we find that some of the faces of the Voronoi region are very small, resulting in sampled interstitial sites very close to each other. Therefore a minimum tolerance of 0.5 Å is used while determining symmetry inequivalent sites. The procedure is described in Fig. 2(b), with ZnO as an example structure.

2.2. Perform defect calculations

As summarized in Fig. 2(a), the workflow starts with fully relaxing (volume, cell shape and ionic positions) the bulk primitive cell. Dielectric constant, and GW calculations are performed on the relaxed primitive cell. Point defects are then created in the bulk supercell followed by relaxation (only ionic positions) of defect structures in multiple charge states. Defect calculations are aimed to use supercell sizes that are large enough to describe individual defects as accurately as possible by minimizing the error due to defect-defect interactions. It has been shown that these interactions are short-ranged and typically occur within the distance of 5–10 Å, from the defect center [7]. Therefore in this work, supercell sizes are chosen such that spacing between defect and its periodic images are greater than 10 Å, and sizes are not too large to make them intractable for DFT calculations. Calculations of interstitial defects are performed in two steps: (1) All candidate interstitials (shown as starting interstitials in Fig. 3) are relaxed in the neutral charge state, (2) followed by relaxation of only unique interstitials (shown as final interstitials in Fig. 3) in multiple charge states. Unique interstitial structures are determined based on: (1) the total energy, (2) space group, and (3) the number of neighboring atoms. The high-throughput DFT calculations are performed using PyLada [23], a Python framework for the organizing and managing high-throughput first-principles calculations. PyLada also offers variety of useful tools for manipulating crystal structures, extracting output from successfully finishes calculations, as well as archiving and analyzing results [25–27]. Finally, the defect formation energies are computed as discussed in the next section.

2.3. Compute defect formation energy

The formation energy of the defect D in the charge state q is calculated as

$$\Delta H_{D,q}(E_F,\mu) = [E_{D,q} - E_H] + \sum_i n_i\mu_i + qE_F + E_{corr} \tag{2}$$

where $E_{D,q}$ and E_H are the total DFT energies of the defect and host supercell, respectively. μ_i is the chemical potential of the atom (host or impurity) of type i added ($n_i < 0$) or removed ($n_i > 0$) from the host supercell to form the defect. E_F is the Fermi energy, and E_{corr} is the term that account for the finite-size corrections, within the supercell approach. A schematic of Eq. (2), representing computation of the defect formation energy from supercell to the dilute limit is shown in Fig. 4.