Simulating charged defects at database scale ⊕

Jimmy-Xuan Shen ✉ 🔟 ; Lars F. Voss 🔟 ; Joel B. Varley 🔟

Check for updates

View Online   Export Citation

Journal of Applied Physics

AIP Publishing

Simulating charged defects at database scale   EP

View Online   Export Citation   CrossMark

Jimmy-Xuan Shen,[1,2,a)] (ID) Lars F. Voss,[3] (ID) and Joel B. Varley[1,2] (ID)

AFFILIATIONS

[1]Materials Science Division, Lawrence Livermore National Laboratory, Livermore, California 94550, USA
[2]Laboratory for Energy Applications for the Future, Lawrence Livermore National Laboratory, Livermore, California 94550, USA
[3]Materials Engineering Division, Lawrence Livermore National Laboratory, Livermore, California 94550, USA

Note: This paper is part of the special topic, Defects in Semiconductors 2024.
[a)]**Author to whom correspondence should be addressed:** shen9@llnl.gov

ABSTRACT

Point defects have a strong influence on the physical properties of materials, often dominating the electronic and optical behavior in semiconductors and insulators. The simulation and analysis of point defects is, therefore, crucial for understanding the growth and operation of materials, especially for optoelectronics applications. In this work, we present a general-purpose Python framework for the analysis of point defects in crystalline materials as well as a generalized workflow for their treatment with high-throughput simulations. The distinguishing feature of our approach is an emphasis on a unique, unit cell, structure-only, definition of point defects which decouples the defect definition, and the specific supercell representation used to simulate the defect. This allows the results of first-principles calculations to be aggregated into a database without extensive provenance information and is a crucial step in building a persistent database of point defects that can grow over time, a key component toward realizing the idea of a "defect genome" that can yield more complex relationships governing the behavior of defects in materials. We demonstrate several examples of the approach for three technologically relevant materials and highlight current pitfalls that must be considered when employing these methodologies as well as their potential solutions.

I. INTRODUCTION

Understanding the physical characteristics of point defects (including dopants) in semiconductors and insulators is a crucial analysis tool in materials and semiconductor research since a small number of defects and dopants can significantly change the electrical and optical properties of a given material.[1] Since defects are usually formed during the initial growth process, defect properties are difficult to isolate and experiment on. Hence, first-principles calculations are often used to analyze these properties. Specifically, calculations using density-functional theory (DFT)[2,3] have been widely adopted as the standard method. The analysis of point defects from first principles usually begins with an analysis of the defect thermodynamics by calculating the formation energy of the defect at a given set of chemical conditions. These formation energies can elucidate which defects are likely to form and different growth conditions and provide information on the doping behavior of these defects.[3] Recently, more advanced analysis capabilities of quantum recombination rates, for radiative,[4] and non-radiative

recombination,[5,6] have been developed to help understand the interactions of electrons and photons around point defects.

The computational cost for performing first-principles defect calculations is unusually high due to a combination of factors. First, the periodic simulation cell containing the defect must be large enough to avoid spurious interactions between the atomic relaxation of the defect and its periodic images. Second, since the defect represents a localized system embedded in a periodic crystal, the system must be simulated using a level of theory that can accurately describe the electronic structure of both the defect and the host crystal. This usually means performing hybrid functional calculations in a plane wave basis set, which is notoriously 10 to 100 times more expensive than traditional local or semi-local functional calculations.

Recently, automated defect analysis tools have been developed to facilitate the simulation and analysis of defects in a high-throughput manner. Notable among these frameworks are pyCDT,[7] pyDefect,[8] AiiDA-defects,[42] doped,[9] and DASP.[10] pyCDT, pyDefect, AiiDA-defects and doped are Python

packages aimed at automating the thermodynamic analysis of defects, while DASP also includes the simulation of various quantum recombination rates.

While these tools accomplish their state goal of automating defect analysis, their adoption by the broader defects research community has been limited and building a large-scale persistent defects database is still out of reach. This is a key limitation towards the realization of a curated "Defect Genome" that can be refined over time and complement the more general Materials Genome Initiative for accelerating materials engineering and discovery.[11] We believe that significant improvements can be made in three areas that will lead to broader adoption: (1) improved integration with existing workflow automation and data aggregation tools, (2) conceptual separation between the code for standard defect analysis and high-throughput calculations, and (3) simulation-independent definition of defects which allows for calculations to be grouped in a database without explicit provenance information. In this work, we present the advances we have made in these areas and demonstrate how a persistent database of defect properties can be built.

The core software driving our defect analysis is pymatgen-analysis-defects,[12] a part of the pymatgen[13] suite materials informatics. To address (1) and (2), we developed pymatgen-analysis-defects to be a general-purpose tool for the analysis of defects in crystalline materials and delegated all of the code for high-throughput workflow automation to the atomate2 package built upon the jobflow framework for automating general high-performance computing tasks.[14] This allows us to codify common defect analysis tasks within pymatgen-analysis-defects and provide a consistent interface for users to perform these tasks without the added complication of high-throughput calculations. For high-throughput calculations, we have developed a generalized defect workflow as part of the atomate2 package, which is capable of orchestrating thousands of DFT defect calculations and organizing their results in a standardized format.

Finally, for (3), we formalized a structure-only definition for point defects which allows us to group calculation results without explicit provenance information. A common critique of high-throughput defect calculations is that structure minimization algorithms can miss the global energy minimum, either due to a lack of local symmetry breaking[9] or not finding the correct spin multiplicity.[15] Having a persistent definition that is independent of the simulation details allows us to revisit previous calculations and find lower energy configurations by using different calculation settings or initial conditions. This allows us to update any database incrementally as more calculations are performed, such as those that increase the accuracy or sample additional structures. Over time, users can improve both the breadth of their defects databases—by including calculations for additional charge state and the quality of data—by finding relaxations to local or global energy minima that were previously missed by the automated workflow.

II. BACKGROUND AND METHODS

The thermodynamic properties of a point defect (X) are controlled by the defect formation energy ($E^f[X^q]$), which measures the energy cost of creating a defect with a charge state q at a given set of chemical conditions. The formation energy is determined by the

chemical potentials of everything that goes into forming the defect[3,7]

$$E^f[X^q] = E_{\text{tot}}[X^q] - E_{\text{tot}}[\text{bulk}] + \sum_i n_i\mu_i + qE_{\text{F}} + \Delta^q, \quad (1)$$

where $E_{\text{tot}}[X^q]$ is the total energy of the defect simulation supercell; $E_{\text{tot}}[\text{bulk}]$ is the total energy of the bulk cell of the same size; n_i is the change in the number of atoms of type i required to form the defect, while μ_i is the chemical potential of atom type i; E_{F} is the chemical potential of the electron (often called the Fermi level); and Δ^q is the finite-size correction required to account for the fact that we are simulating a charged defect in a finite periodic cell.

While all of the different chemical potentials are interconnected, it is conceptually easier to separate them into two categories: the chemical potential of the atoms and the chemical potential of electrons. The chemical potential of electrons (often called the Fermi level) accounts for how all of the external conditions, including the presence of other defects, affect the energy costs of adding or removing an electron from the defect system. Since the electrons in the system can be manipulated after the defect is formed, the Fermi level is often considered a free variable and is shown as the x-axis whenever $E^f[X^q]$ is plotted for a given set of the other chemical potentials. The chemical potentials of the different atoms added to or removed from the defect are less dynamic and assumed to be fixed after the defect is formed. So, the formation energy of each charge state of a defect is typically shown as a linear function of E_{F} with a slope of q with a y-intercept determined by the chemical environment during crystal growth.

The core goal of any defect analysis code is to calculate the formation energy of a given defect in the different charge states. However, since $E_{\text{tot}}[X^q]$ is calculated by creating a point defect in a periodic supercell and Δ^q presumably eliminates the finite size effects, we can have many different supercell calculations that all correspond to the same defect charge state. Traditionally, users will have to track the provenance of the defect and charge state for all the calculations. This can quickly become untenable as the number of defects and charge states increases. Additionally, since the most stable spin and atomic configuration of a defect might not be found by the first structure relaxation, it is often necessary to revisit defect calculations to improve upon the data over time. This necessitates a way for us to compare and group defects without explicit provenance information. Toward that end, we have developed a structure-only definition of point defects that will allow persistent, incrementally growing databases of defect simulation results to be constructed (cf. Sec. II A). To compute the different atomic chemical potentials (μ_i), we combine explicit DFT calculations of the pure elemental phases and experimentally corrected values of the formation enthalpy to arrive at a more accurate measure of the elemental chemical potentials with fewer calculations.

For the calculation of the finite-size correction term (Δ^q), we use the finite-size correction method of Freysoldt, Neugebauer, and Van de Walle (FNV).[16] Previous implementations[7,16] of this method required the position of the defect to be supplied by the user. Since we require our data to be provenance-agnostic, we have developed a method to calculate the position of a point defect without prior knowledge of how the defect was defined (cf. Sec. II C). This eliminates a persistent point of user

intervention in the workflow of defect analysis and allows for more software-driven high-throughput approaches.

The combinations of these advances allow us to arrive at an analysis framework to define, simulate, and analyze defects in a fully automated manner. All of the code responsible for the definition and analysis of point defects are available via the pymatgen-analysis-defects add-on package to pymatgen.[13] The automation workflow is distributed independently as part of the atomate2 package which allows us to orchestrate and organize thousands of quantum chemistry calculations (cf. II D). We will detail these advancements in the remainder of this section.

A. Structure-only definition of defect

For most DFT calculations, the physical system you are simulating is solely defined by the atomic positions and the lattice vectors of the unit cell. This allows large databases of DFT calculations to be built by simply storing the atomic structures and aggregating all of the calculations related to a given structure into a single entry. While the unit cell is not unique, modern symmetry analysis tools allow different representations of the same periodic structure to be matched to each other so it is trivial to group calculations that represent the same physical system. However, when simulating point defects, we are using an arbitrary supercell to simulate the defect in the dilute limit, so multiple symmetry-distinct supercell structures can all correspond to the same point defect. This inherently breaks the one-to-one mapping between structures and calculations, and thus the structure matching along is not enough facilitate the building of a persistent database of defect calculations. A core feature of pymatgen-analysis-defects is the ability to define defects in a structure-only manner, which allows us to group calculations for the same defect without explicit provenance information. Ultimately, this definition frees us from the need to explicitly track the relationship and dependencies between sets of defect calculations and makes database building significantly easier.

In pymatgen-analysis-defects, a defect is defined by the combination of

- The type of defect: whether it is a vacancy, interstitial, or substitutional defect.
- A unique primitive unit cell of the bulk material.
- The fractional position of the defect in the unit cell.

This definition allows us to create a representative structure ($\mathcal{S}_X$) of the defect X in the unit cell. As long as the primitive cell is fixed, two defects X and Y are considered to be the same if and only if $\mathcal{S}_X$ and $\mathcal{S}_Y$ can be structurally matched to each other, i.e., they are equivalent under the symmetry operations of the primitive cell. While there might be multiple equivalent sites in a primitive cell for a point defect to form, the structures representing these different versions will be symmetrically equivalent to each other. As examples, all of the symmetry-distinct native defects of wurtzite GaN are shown in Fig. 1. The vacancy and antisite defects are trivially constructed, however, generating interstitial defects has always been challenging. Following the method from Ref. 17, the symmetry-distinct (as defined using Spglib[18]) local minima in the charge density were identified as possible interstitial insertion sites, which resulted in two symmetry-distinct versions of the Ga and N interstitial defects. Note

that this method and be easily combined with charge density application programming interface[19] from the Materials to quickly generate defect structures without additional DFT calculations.

These representative defect structures can be used to define additional quantities such as predicted oxidation states of the elements and degeneracy factors, which are required to calculate various thermodynamic quantities related to point defects.[3] Additionally, once a defect has been created, we can generate a nearly cubic supercell structure containing that point defect to minimize the effects of periodic images. The resulting calculations can be associated with the original defect so the thermodynamic quantities from the DFT calculations amount to one representation of a concrete concept such as "energy of V_{Ga} in the $q = -1$ charge state." For the 4-atom primitive cell of GaN, we can create a 128-atom supercell that has orthogonal lattice vectors that differ in length by less than 10%. For the remainder of this work, explicit supercell defect DFT calculations will be performed using this GaN supercell structure.

B. Chemical potential corrections based on experimental data

The chemical potential is a measure of the thermodynamic cost of adding or removing an atom from the defect when the formations of different competing phases are considered. If the formation of the defect is only competing with the pure elemental phases, then the chemical potential is simply the energy per atom from the DFT calculation of the elemental phase (μ_i^{DFT}). However, competing compounds with lower formation enthalpy can lower the chemical potential ($\mu_i = \mu_i^{DFT} + \Delta\mu_i$) since they create a more stable environment for the defect. For a set of atomic species (indexed by i), and a set of competing compounds (indexed by α) these correction terms must satisfy the following constraints:

$$\Delta\mu_i \leq 0 \qquad \forall \text{ elements } i,$$
$$\sum_\alpha n_i^\alpha \Delta\mu_i < \Delta H_f(\alpha) \qquad \forall \text{ compounds } \alpha,$$
$$\sum_\alpha n_i^{bulk} \Delta\mu_i = \Delta H_f(bulk), \qquad (2)$$

where n_i^α is the number of atoms of species i in a formula unit of compound α, and the $\Delta H_f(\alpha)$'s are the formation enthalpies of one formula unit of the compound α. Equation (2) defines a manifold in chemical-potential space where you are constrained to the facet representing the host material but bound on different sides by the formation of competing phases. Since the total energy from DFT can change under different calculation settings, we are usually forced to calculate all of the formation enthalpies using the same level of theory. However, since the formation energy data from the Materials Project is specifically corrected to fit experimental formation enthalpies,[20] we can use the formation enthalpy data from the Materials Project to calculate the chemical potential corrections ($\Delta\mu_i$) only and expect similar if not better results. Hence, the chemical potential in Eq. (1) can be expressed as

$$\mu_i = \mu_i^{DFT} + \Delta\mu_i^{MP}, \qquad (3)$$

where we combine the explicitly calculated energy of the elemental phase in a given level of theory with the formation enthalpy correction due to competing phases from the Materials Project. This approach is
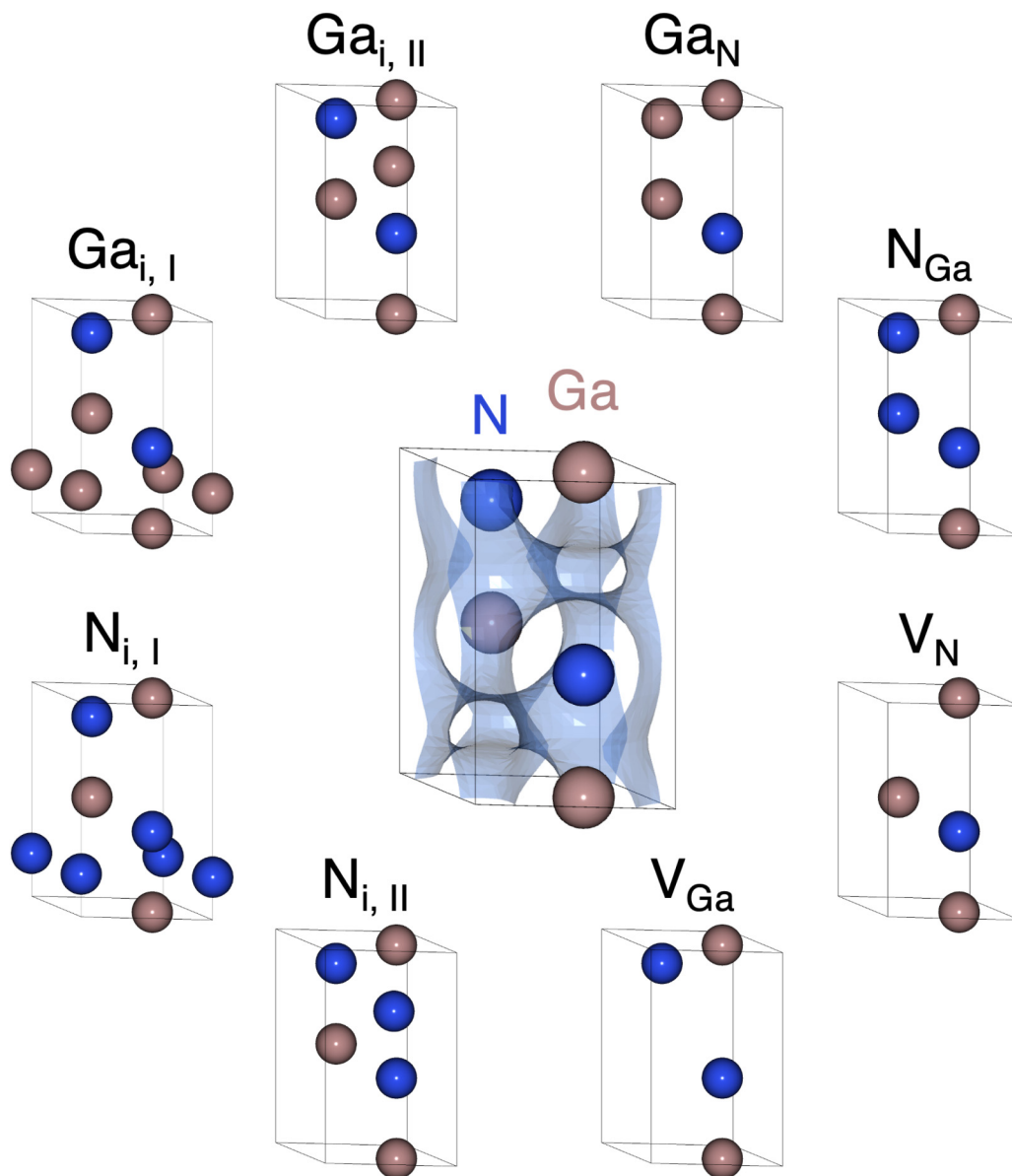
20 April 2024 06:39:58

FIG. 1. Defect structures for native point defects in GaN, note that two symmetry distinct initial interstitial sites were identified.

distinct from other approaches based on correcting inaccuracies in chemical potential references, such as fitted elemental-phase reference energies (FEREs),[21] which instead corrects the μ_i^{DFT} terms based on a consistent calculation framework and thermodynamic databases from which to derive the corrections. From a database-building perspective, this simultaneously provides a more accurate prediction of relative stability between the competing phases (since we are using experimentally-fitted formation enthalpies) and allows users to build a database using only the DFT energy from the elemental phases alone, which significantly reduces the computational cost.

C. Identify defect position with no prior knowledge

The total energy of the defective supercell $[E^f[X^q]$ in Eq. (1)] refers to the energy of the defect simulation supercell after atomic relaxation. Since atomic relaxation is often unpredictable, it is difficult to know the exact position of the defect in the supercell afterward. This is especially problematic when we attempt to calculate the finite-size charge correction for native defects in a supercell since the position of the defect is an input that helps determine the potential alignment contribution.[16] Usually, one can arrive at a

decent estimate of the defect position by keeping track of the position during the defect-creating processes. But for cases where the relaxation of the atomic positions is severe, such as the case of split interstitials and split vacancies, the relaxed defect position will differ significantly from the defect creation site. Additionally, as we aim to build databases without explicit provenance information, we need to find a way to identify the position of the defect in the supercell after structure relaxation without any prior knowledge of the defect creation site. In pymatgen-analysis-defects, we included a method to identify the position of the defect in the supercell after structure relaxation by calculating a site-specific distortion field based on SOAP vectors.[22,23] For each site in the relaxed supercell, we calculate the SOAP vectors ($\boldsymbol{v}_\sigma^{\text{defect}}$) for each site σ and compare them with the SOAP vectors of the sites in the pristine bulk structure ($\boldsymbol{v}_{\sigma'}^{\text{bulk}}$). The distortion on site σ is defined as

$$\delta_\sigma = 1 - \max_{\sigma'}\left\{ \frac{\boldsymbol{v}_\sigma^{\text{defect}} \cdot \boldsymbol{v}_{\sigma'}^{\text{bulk}}}{\left|\boldsymbol{v}_\sigma^{\text{defect}}\right|\left|\boldsymbol{v}_{\sigma'}^{\text{bulk}}\right|} \right\}. \tag{4}$$

To identify the set of sites with the largest distortions, we first sort the distortion values (δ_σ) in descending order and cut off the descending series at the largest drop in distortion value. The position of the defect is then given by taking the average position of the most distorted sites, weighted by their distortion value.

Using this site finder algorithm, we can calculate the FNV finite-size correction with only the electrostatic potentials from the DFT calculations and the dielectric constant of the material. The alignment-like term in the original Freysoldt correction paper requires the short-range potential ($V_{q/0}^{\text{sr}}$) to be shifted by a constant value (C) so that the potential is effectively zero far away from the defect. Figure 2 shows the planar-averaged values of the different potentials involved in calculating the FNV charge-state correction. The difference in electrostatic potentials between the defect and pristine cells (black curve) and the short-range potential (green curve) are both directly computed from the files storing the electrostatic potential information (e.g., LOCPOT files for VASP calculations). The position of the defect was automatically determined using the finder algorithm and can be validated by the fact that the potential difference and short-range potential profile both peak at the origin.

This method removed the usual need for user input from previous methods[7,16] and allow for the entire process of defect formation energy calculation to be codified within our Python framework. This makes the investigation of singular defects much more approachable and significantly reduces the book-keeping complexity incurred in other high-throughput defect analysis approaches. Additionally, the ability to easily identify the positions of native vacancies and interstitials in structures using only the atomic positions is likely to have uses beyond the present context.

D. atomate2 workflow for defect formation energies

The release of atomate[24] and aiida[25] marked a major milestone in the open-source materials workflow development, as they codified many standard materials science workflow under a well-defined automation framework. However, complex workflows such as defect formation energy calculation were difficult to develop due
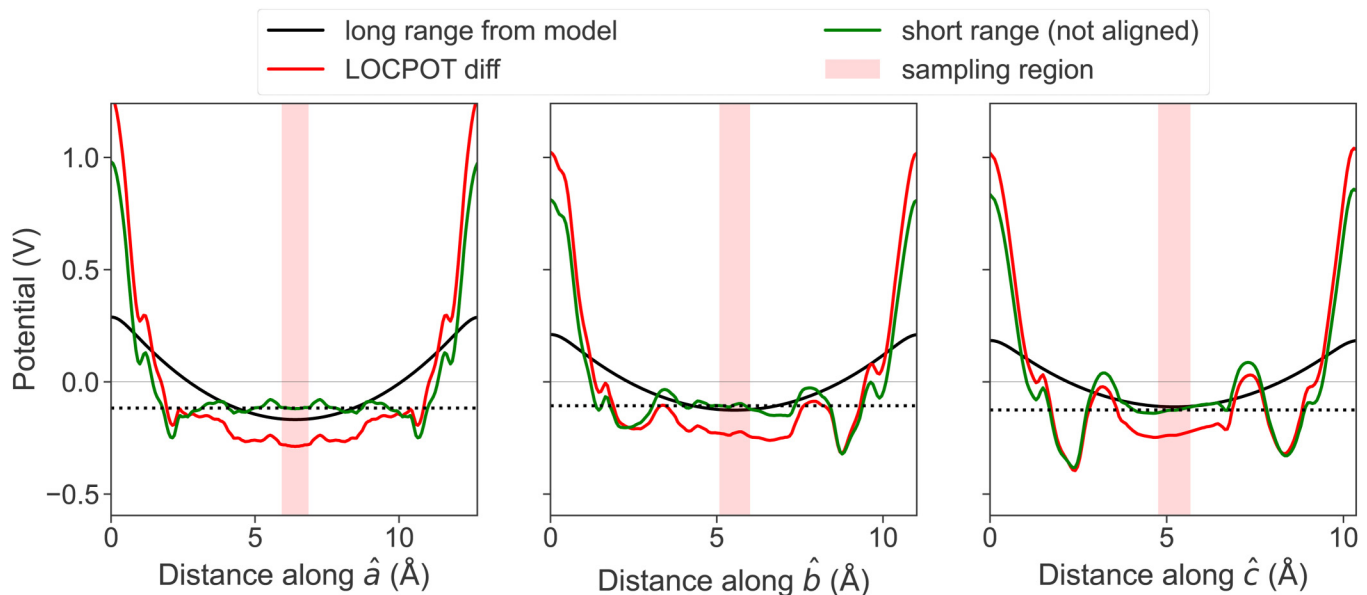
FIG. 2. Planar-averaged potentials used during the calculations of the FNV charge-state correction for the V_{Ga} defect in GaN. The long-range (black) term is computed from a Gaussian model, the potential difference (red) and short-range terms are both shifted so the origin is at the defect positions identified by the finder algorithm. The sampling region (shaded red), over which the C offset (dashed) is computed in a small region far from the origin in each crystal direction.

to the fundamental limitations such as processing of the volumetric data and the ability to dynamically create additional DFT calculations based on the results of previous calculations. atomate2 directly addresses these challenges and allows us to define dynamic and composable workflows that can also natively handle the storage of large volumetric data. With these advances, we were able to develop a flexible and composable high-throughput defect simulation framework that is capable of handling a variety of different simulation needs.

The flow chart for the workflow to calculate the defect formation energy of a single defect is shown in Fig. 3. The two primary inputs for the workflow are the uniquely defined Defect object (as discussed in Sec. II A) and the Relaxer construct that contains all of the specific calculation settings for the relaxations procedure. Because we have abstracted away the computational details of the charge state relaxation using the Relaxer, any functional support by VASP[26] or even entirely different DFT simulation codes (such as CP2K[27]) can be used with minimal modification to the code.

In the workflow for a single defect, we first create a supercell of the bulk material and relax the lattice parameters and atomic
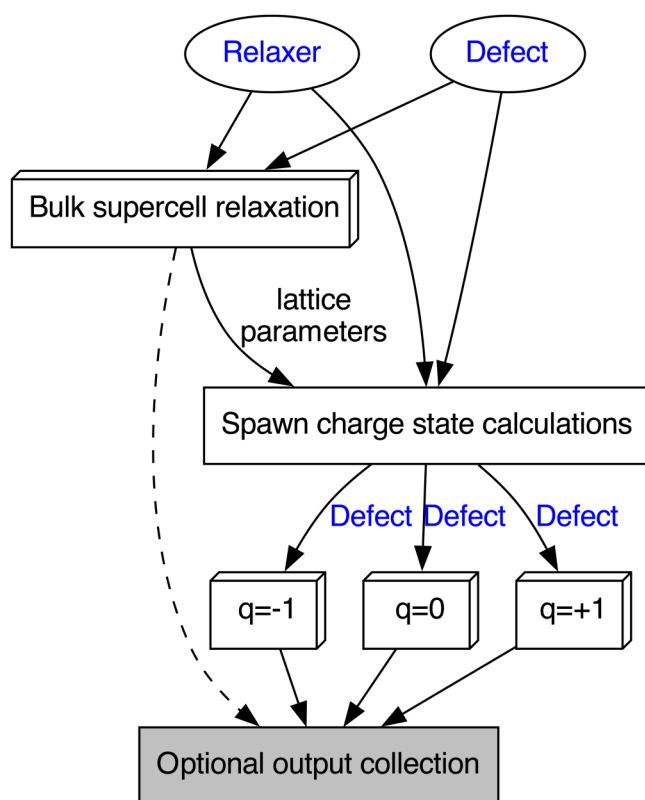

FIG. 3. Flow chart for the workflow to calculate the defect formation energy. The Relaxer and Defect objects are the required inputs. A single bulk supercell calculation is required to obtain the electrostatic potential (e.g., a LOCPOT file for VASP calculations) data and to determine the supercell lattice parameters. The outputs from the different charge state supercell calculations can be combined with the bulk calculation to compute the defect formation energy.

positions. Then, we create a defect supercell with the same lattice parameters as the bulk supercell and relax only the atomic positions of each charge state. The charge states in this case, are derived from a combination of the allowed formal oxidation states of the species in the bulk formula, and the abundances of specific charge states for these species in the ICSD.[28] Once these individual charge state relaxations are completed, their raw results are stored in a database along with a serialized version of the original defect object used to generate these calculations. This allows the final data aggregation step to only examine the Defect objects and the individual calculation settings to quickly find the set of calculations corresponding to charge states of a particular defect simulated using a given level of theory. The final step is to combine the computed energies and electrostatic potential data from the bulk calculation and different charge state calculations to arrive at a prediction of the formation energy of the defect. Since we have designed our framework specifically to eliminate the need to track data provenance, this final step can be performed independently from the simulation process. This becomes more important as a database scales and bookkeeping for the provenance of specific calculations becomes more difficult and mistake-prone.

The composable nature of atomate2 allows us to connect many defect workflows to avoid repeating the *Bulk supercell relaxation* step. We can perform this step once and initialize the *Spawn charge state calculations* step without any additional DFT calculations.

III. RESULTS & DISCUSSIONS

Using the framework described in Sec. II, we have calculated the formation energies of all of the native defects in GaN at different levels of theory and under different atomic chemical conditions. The formation energy diagrams for native defects of GaN, computed using the screened hybrid functional from Heyd–Scuseria–Ernzerhof[29] and PBEsol functionals, are shown in Fig. 4. We elected not to fit the exact-exchange fraction to the bandgap and used a default of 25% to make the data directly comparable between material systems. Nonetheless, the results of our HSE06 calculations match well with previously reported results with a different exact exchange that better represents the experimental bandgap of GaN[30] with one notable exception of the N_{Ga} defect. For N_{Ga}, the formation energy of the $q = 0$ charge state in HSE06 was too high to be stable relative to the $q = -1$ and $q = +1$ charge states and the estimated charge state range did not agree with the charge states found by Lyons *et al.*[30] We use this example to illustrate the first shortcoming in typical high-throughput defect approaches that generate structures that may adopt insufficiently perturbed initial structures or an incomplete set of defect charge states (based on the constituents' oxidation states) that are explored. The first problem can be fixed by performing the HSE06 calculation with larger distortions to the environment around the defect, which lowered the formation energy of the $q = 0$ charge state by 0.58 eV and stabilized the neutral charge state in *n*-type conditions. The second problem is more subtle and caused by the unusual oxidation behavior of this particular defect. Since Ga is found in the $+3$ oxidation state and N is found in the -3 oxidation state the relevant charge states for the N_{Ga} defect are expected to range from
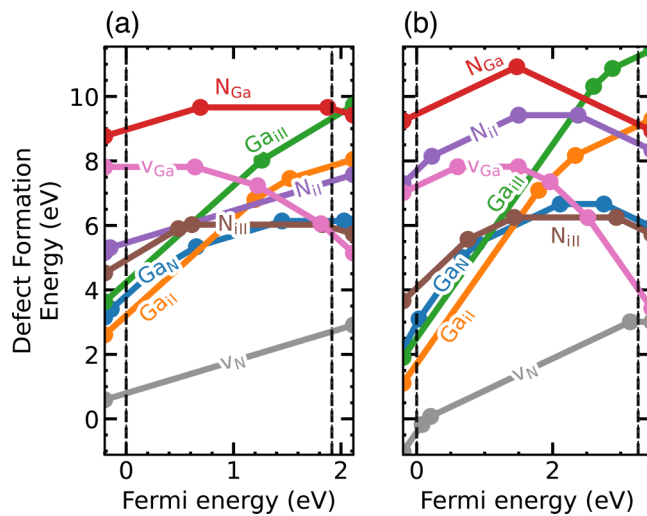
FIG. 4. Formation energy diagrams for native defects in GaN computed using PBEsol (a) and HSE06 (b). Both sets for formation energy diagrams are calculated using Ga-rich chemical conditions.

-6 to 0, however, Lyons *et al.*[30] found multiple positive charge states for this defect to be stable. This represents one example of how "standard" types of point defects may be interpreted as defect complexes for particular oxidation states, e.g., V_{Ga}–N_i complexes that can adopt higher positive charge states for Fermi levels closer to the VBM. We note that this type of consideration can be factored in *a priori* for the prototypical point defect types considered here. While formal treatment of defect complexes, as well as the integration with informed symmetry-breaking algorithms[9] are fully supported by the pymatgen-analysis-defects framework, they should only be considered on a case-by-case basis due to the increased computational cost. While these issues only affected a single high-formation-energy defect in those studied in GaN, they highlight the general difficulty in automating defect calculations and why our approach to incrementally updating the data is required.

We compared our PBEsol and HSE06 calculations and discovered two general methods to improve the overall speed and reliability of defect calculations. First, we found that the use of a PBEsol calculation to precondition the HSE06 relaxation can significantly reduce the computational cost of the defect calculations. However, we acknowledge that this must be considered with care for certain charge states, as a bias toward charge delocalization from semilocal functionals can lead to relaxed structures in local minima, which may require additional perturbations or relaxation steps to access global minima for a given defect configuration and charge state.[9] Second, we found that constraining the atoms far from the defect to their bulk positions can both reduce the computational cost and improve the reliability of the defect calculations. Both of these approaches have been implemented as optional settings that can be easily toggled on or off by the user.

We recommend performing HSE06 defect calculations using a Relaxer that consists of a fast PBEsol relaxation to precondition

the structure and wavefunctions, followed by a slower HSE06 relaxation which typically produces more accurate descriptions of the localized defect states and energies. We again stress that the extent to which the defect structures are relaxed with (semi)local functionals before switching to hybrid functionals is sensitive to the particular oxidation state and symmetry of the defect, with electronic closed-shell defects being much more amenable to extended relaxations with semilocal functionals, such as the V_{Ga}^{-3}. For 8 native defects in GaN, we identified a total of 70 distinct charge state calculations that needed to be performed based on our analysis of the formal oxidation state of the defects. Each electronic optimization step of an HSE06 calculation is usually around 10–200 times more expensive than the corresponding PBEsol calculations so we can simply count to number of HSE06 electronic steps taken to determine the relative computational cost of the different approaches. Using only HSE06 to perform the atomic relaxations, we needed an average of 132.3 electronic steps to complete each charge state calculation. Using PBEsol atomic to precondition the HSE06 relaxation, we needed an average of 84.2 HSE06 electronic steps to complete each charge state calculation. This represents a 36% reduction in the number of electronic steps required with a less than 0.05 eV difference in the final formation energy predictions.

A well-known problem when performing defect calculations is the fact that the perturbation of the crystal caused by the defect can cause another phase transition in the host materials so no matter how large the supercell is, the defect distortion caused by the defect will still cause periodic image interactions. This can be especially problematic for interstitial defects which can cause large distortions in the surrounding lattice. These distortions are often not physically meaningful since the bulk phase is stabilized by finite temperature effects not captured by the ground-state DFT calculations. To avoid this problem, we first calculate the largest sphere that can be inscribed in the supercell and then constrain all atoms outside of this sphere to their bulk positions. This ensures that the lattice distortions caused by the defect cannot interact with itself across periodic boundaries.

Using the same methods for generating and analyzing native defects in GaN, we have also computed the formation energies of the native defects in β-Ga_2O_3, a lower-symmetry, more complicated monoclinic structure that has two symmetry-distinct Ga sites and three symmetry-distinct O sites. This leads to 10 distinct "simple" native defects in Ga_2O_3, excluding interstitials: two each for the V_{Ga} and O_{Ga} defects and three each for the V_O and Ga_O defects. with a number of interstitial sites also possible in the β-gallia structure. Using the same charge-density method as above,[17] we identified five symmetry distinct interstitial sites by analyzing the electronic charge density and they coincide with recently published work on interstitials in Ga_2O_3.[31] Since the number of candidate interstitial sites can be large, we recommend limiting the insertions to the two sites with the lowest average charge density in the surrounding region since that is known to be a good predictor of stability.[32] The formation energy diagrams for native defects in Ga_2O_3 are again computed with two different levels of theoretical treatment, this time with the PBEsol and HSE06 functionals and shown in Fig. 5. As for GaN, we used a standard fraction of exact exchange of 25% for HSE, which is also known to underestimate
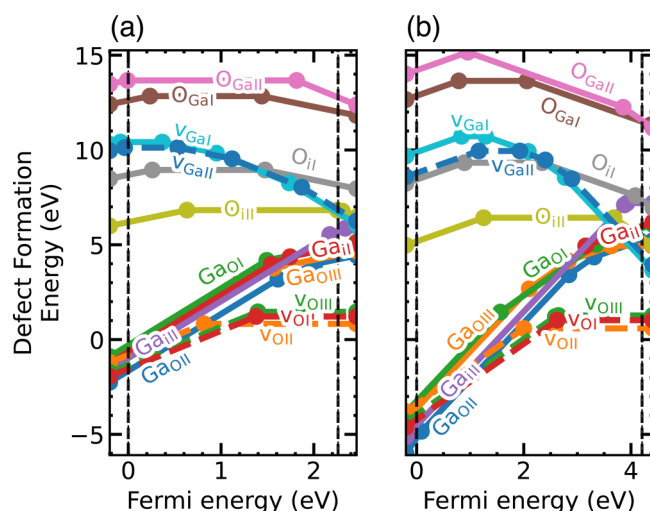
FIG. 5. Formation energy diagrams for native defects in Ga_2O_3, computed using PBEsol (a) and HSE06 (b). Both sets for formation energy diagrams are calculated using Ga-rich chemical conditions.

the bandgap of β-Ga_2O_3. Once again, our fully automated framework largely replicated recently published computational results[33] without any human intervention. We note that the qualitative orderings of the defects are consistent with other results, and summarize the computed transition levels in the supplementary material. However, it has been shown that β-Ga_2O_3 exhibits several highly distorted vacancy configurations (so called split-vacancies) that are lower in energy and that the automated relaxation framework was unable to initially capture.[33–35] This echoes a key current deficiency in point defect workflows similar to the case with GaN, and highlights the importance of efficient databasing and the ability to apply additional structural perturbations that may be outside the typical tolerance of automated symmetry-breaking algorithms.

As a third case study, we also computed the formation energies of the different native defects of cubic $SrTiO_3$ (STO). While the cubic perovskite phase is structurally simple, STO thermodynamically prefers a lower-symmetry structure at 0 K and also exhibits a more complex chemical stability region given the larger number of elemental constituents. The perturbation caused by the defect can lead to spurious transformations to the cubic phase when the defect is simulated in a finite-sized periodic cell. To alleviate this, we define a sphere centered at the defect with the largest radius allowed by the periodic boundary conditions and freeze all atoms outside this sphere. This leads to faster convergence of the structural relaxation and more reliable results.

The atomic structure of STO only has a single symmetry-distinct site for each atomic species, resulting in three distinct vacancies and six distinct antisite defects. Two symmetry-distinct interstitial sites were identified using the electronic charge density, which resulted in another six interstitial native point defects. Since the Sr–Ti–O chemical space is more complex, the benefit of using

thermodynamic databases like the Materials Project to determine the relevant boundaries for the chemical potentials is more apparent. A total of nine distinct compounds are involved in defining the stability region of STO in chemical potential pace, which would usually require nine additional hybrid DFT calculations on top of the calculations for the total energies of the elemental phases. Using the method outlined in Sec. II, we can skip these nine calculations and use the formation enthalpy data from the Materials Project to compute the chemical potential limits. The stability region for STO in chemical potential space is shown in Fig. 6(a), and we have selected the Ti-rich and O-rich growth conditions to represent two limiting conditions for the growth of STO. This contrasts previous studies[36,37] which only considered TiO_2 and SrO as the competing phases and only considered a single interstitial site.

The formation energy diagrams of native defects of STO under O-rich conditions are shown in Figs. 6(b)–6(d) and the ones for Ti-rich conditions are shown in Figs. 6(e)–6(g). Similar to Refs. 36 and 37, we find that the three vacancies V_O, V_{Ti}, and V_{Sr} are the most stable native defects under most conditions. Under O-rich conditions, the +2 charge state of V_O is the most stable defect across most Fermi levels in the band gap. Under Ti-rich conditions, the −2 charge state of V_{Sr} becomes more stable for n-type conditions, and the −4 charge states of V_{Ti} become stable just below the conduction band. The relevant defects, charge states, and formation energies of these defects change under different chemical conditions are all consistent with previous studies[36,37] with errors typical for defect calculations.[38–40]

As we have seen from the examples above using GaN, Ga_2O_3, and STO, our fully automated framework can reproduce previous computational results without any human intervention. Instead of a static database of the results at the end, our strategy results in a dynamic database that can be incrementally improved over time as more accurate calculations are performed. In each example above, we were able to identify a set of defects that are likely to be relevant under a given set of chemical conditions and have a reasonable estimate of the charge state behavior of these defects.

A limitation of any high-throughput approach is the trade-off between generality and accuracy. Here, we have elected to use a single mixing parameter of 25% for all calculations, which is a choice that simplifies the comparison of energetics within a given level of theory and computational parameters. However, a fixed mixing parameter prevents us from using the common strategy for tuning the mixing parameter to match the experimental bandgap, a typical choice for more direct comparisons of calculated defect transition levels with experiments. Choices of "incorrect" exact exchange will result in some inaccuracies in the formation energies and defect transition levels that are largely predictable and quantifiable given that localization is qualitatively consistent between the calculations and unlikely to affect the overall trends in the defect behavior.[41] This re-emphasizes the need to separate the responsibilities of the defect analysis and high-throughput defect simulation since high-throughput defect analysis is often more concerned with the relative stability of the defects and not the exact thermodynamic properties of a specific defect. The data shown in Figs. 4–6 are available as tables in the supplementary material.
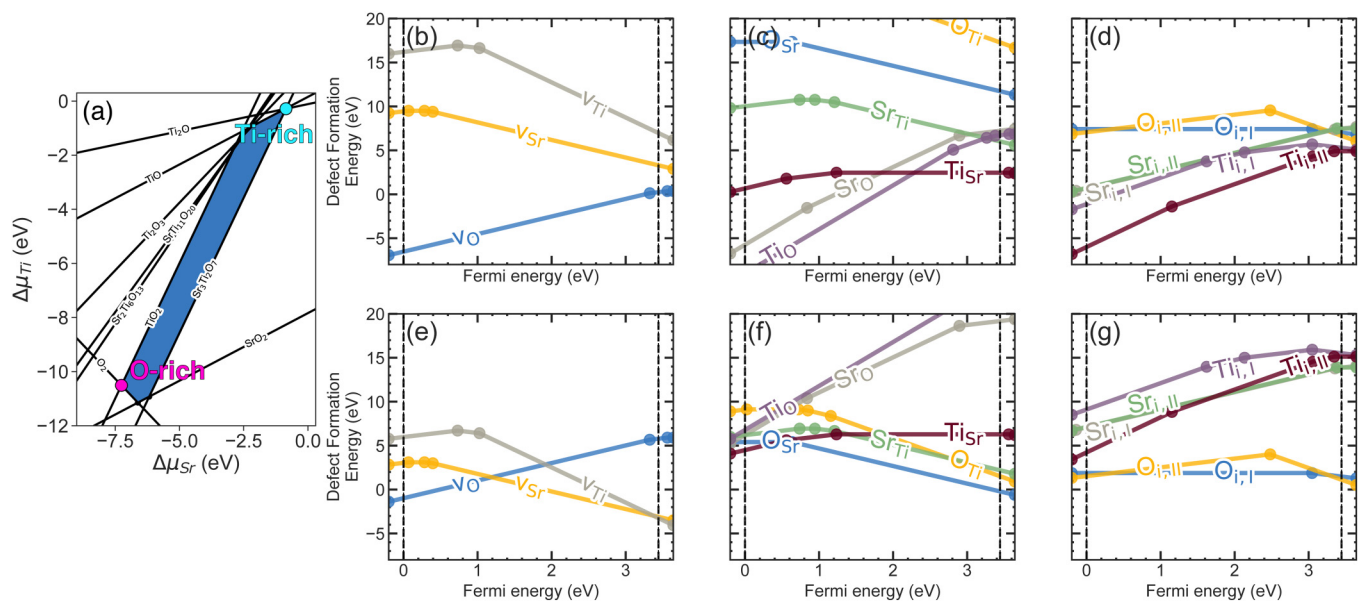
FIG. 6. (a) The stability region of STO in the (Sr, Ti) chemical potential space, the Ti-rich (teal), and O-rich (magenta) growth conditions are indicated. (b)–(d) The formation energy diagrams of native point defects of STO under Ti-rich conditions. (e)–(g) The formation energy diagrams of native point defects of STO under O-rich conditions.

IV. CONCLUSION

We have demonstrated a general-purpose codebase (pymatgen-analysis-defects) for the analysis of point defects in crystalline materials. Our code enforces a strong connection between a symmetry-distinct defect structure and core concepts in defect analysis, allowing for further extension of the code base in an object-oriented fashion. Although not discussed in the manuscript, we have also implemented methods for configuration-coordinate diagram analysis as well as tools for simulating radiative and non-radiative recombination at defect centers. This framework codifies some of the more complex computational workflows in materials informatics and the strong integration with existing analysis and automation software allows for more consistent development and maintenance by the community as the core pymatgen codebase has demonstrated over the past decade.

Since the code is designed for minimal user intervention, we were able to integrate this with existing workflow automation tools to build a fully automated high-throughput defect analysis framework. We have demonstrated that a fully automated approach can largely reproduce previously published computational results without any human intervention and identified current drawbacks with the current implementation. Specifically, we find that under-sampling of global and local minima structures of particular oxidation states, particularly in the case of defect complexes or other defects that exhibit large structural distortions, may be missed in initial searches but can be readily revisited with our emphasis on a robust database-centered representation of point defects. These issues are currently being resolved in future versions that will handle symmetry-breaking and defect complexes in a more

rigorous manner. With more computational resources, it is now possible to build a persistent database of defect properties that can be improved over time as more accurate calculations are performed, although that is beyond the scope of the present work.

SUPPLEMENTARY MATERIAL

The transition levels for GaN, Ga_2O_3, and $SrTiO_3$ are presented as tables in the supplementary material.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Jimmy-Xuan Shen: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Resources (supporting); Software (lead); Visualization (lead); Writing – original draft (lead). **Lars F. Voss:** Funding acquisition (equal); Writing – review & editing (supporting). **Joel B. Varley:** Funding acquisition (equal); Methodology (equal); Validation

(equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material. The code used to generate the data is released under the modified BSD license and is available in GitHub at https://github.com/materialsproject/pymatgen-analysis-defects, Ref. 43.

REFERENCES

[1] H. McCluskey, *Dopants and Defects in Semiconductors* (Taylor & Francis, Andover, 2012).

[2] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," Phys. Rev. **140**, A1133–A1138 (1965).

[3] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, "First-principles calculations for point defects in solids," Rev. Mod. Phys. **86**, 253–305 (2014).

[4] C. E. Dreyer, A. Alkauskas, J. L. Lyons, and C. G. Van de Walle, "Radiative capture rates at deep defects from electronic structure calculations," Phys. Rev. B **102**, 085305 (2020).

[5] D. Wickramaratne, J.-X. Shen, C. E. Dreyer, M. Engel, M. Marsman, G. Kresse, S. Marcinkevičius, A. Alkauskas, and C. G. Van de Walle, "Iron as a source of efficient Shockley-Read-Hall recombination in GaN," Appl. Phys. Lett. **109**, 162107 (2016).

[6] A. Alkauskas, Q. Yan, and C. G. Van de Walle, "First-principles theory of non-radiative carrier capture via multiphonon emission," Phys. Rev. B **90**, 075202 (2014).

[7] D. Broberg, B. Medasani, N. E. R. Zimmermann, G. Yu, A. Canning, M. Haranczyk, M. Asta, and G. Hautier, "Pycdt: A python toolkit for modeling point defects in semiconductors and insulators," Comput. Phys. Commun. **226**, 165–179 (2018).

[8] Y. Kumagai, N. Tsunoda, A. Takahashi, and F. Oba, "Insights into oxygen vacancies from high-throughput first-principles calculations," Phys. Rev. Mater. **5**, 123803 (2021).

[9] I. Mosquera-Lois, S. R. Kavanagh, A. Walsh, and D. O. Scanlon, "Identifying the ground state structures of point defects in solids," npj Comput. Mater. **9**, 1–11 (2023).

[10] M. Huang, Z. Zheng, Z. Dai, X. Guo, S. Wang, L. Jiang, J. Wei, and S. Chen, "Dasp: Defect and dopant ab-initio simulation package," J. Semicond. **43**, 42101 (2022).

[11] Q. Yan, S. Kar, S. Chowdhury, and A. Bansil, "The case for a defect genome initiative," Adv. Mater. **36**, 2303098 (2024).

[12] J.-X. Shen and J. Varley, "pymatgen-analysis-defects: A python package for analyzing point defects in crystalline materials," J. Open Source Soft. **9**, 5941 (2024).

[13] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, "Python materials genomics (pymatgen): A robust, open-source python library for materials analysis," Comput. Mater. Sci. **68**, 314–319 (2013).

[14] A. S. Rosen, M. Gallant, J. George, J. Riebesell, H. Sahasrabuddhe, J.-X. Shen, M. Wen, M. L. Evans, G. Petretto, D. Waroquiers, G.-M. Rignanese, K. A. Persson, A. Jain, and A. M. Ganose, "Jobflow: Computational workflows made simple," J. Open Source Soft. **9**, 5995 (2024).

[15] Y. Xiong, C. Bourgois, N. Sheremetyeva, W. Chen, D. Dahliah, H. Song, S. M. Griffin, A. Sipahigil, and G. Hautier, "High-throughput identification of spin-photon interfaces in silicon," arXiv:2303.01594 (2023).

[16] C. Freysoldt, J. Neugebauer, and C. V. de Walle, "Fully ab initio finite-size corrections for charged-defect supercell calculations," Phys. Rev. Lett. **102**, 016402 (2009).

[17] J.-X. Shen, M. Horton, and K. A. Persson, "A charge-density-based general cation insertion algorithm for generating new Li-ion cathode materials," npj Comput. Mater. **6**, 161 (2020).

[18] A. Togo and I. Tanaka, "Spglib: a software library for crystal symmetry search," arXiv:1808.01590 (2018).

[19] J.-X. Shen, J. M. Munro, M. K. Horton, P. Huck, S. Dwaraknath, and K. A. Persson, "A representation-independent electronic charge density database for crystalline materials," Sci. Data **9**, 1–7 (2022).

[20] A. Wang, R. Kingsbury, M. McDermott, M. Horton, A. Jain, S. P. Ong, S. Dwaraknath, and K. A. Persson, "A framework for quantifying uncertainty in DFT energy corrections," Sci. Rep. **11**, 15496 (2021).

[21] V. Stevanović, S. Lany, X. Zhang, and A. Zunger, "Correcting density functional theory for accurate predictions of compound enthalpies of formation: Fitted elemental-phase reference energies," Phys. Rev. B **85**, 115104 (2012).

[22] A. P. Bartók, R. Kondor, and G. Csányi, "On representing chemical environments," Phys. Rev. B **87**, 184115 (2013).

[23] S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, "Comparing molecules and solids across structural and alchemical space," Phys. Chem. Chem. Phys. **18**, 13754–13769 (2016).

[24] K. Mathew, J. H. Montoya, A. Faghaninia, S. Dwarakanath, M. Aykol, H. Tang, I. heng Chu, T. Smidt, B. Bocklund, M. Horton, J. Dagdelen, B. Wood, Z.-K. Liu, J. Neaton, S. P. Ong, K. Persson, and A. Jain, "Atomate: A high-level interface to generate, execute, and analyze computational materials science workflows," Comput. Mater. Sci. **139**, 140–152 (2017).

[25] S. P. Huber, S. Zoupanos, M. Uhrin, L. Talirz, L. Kahle, R. Häuselmann, D. Gresch, T. Müller, A. V. Yakutovich, C. W. Andersen, F. F. Ramirez, C. S. Adorf, F. Gargiulo, S. Kumbhar, E. Passaro, C. Johnston, A. Merkys, A. Cepellotti, N. Mounet, N. Marzari, B. Kozinsky, and G. Pizzi, "AiiDA 1.0, a scalable computational infrastructure for automated reproducible workflows and data provenance," Sci. Data **7**, 1–18 (2020).

[26] G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set," Phys. Rev. B **54**, 11169–11186 (1996).

[27] T. D. Kühne, M. Iannuzzi, M. D. Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, and J. Hutter, "Cp2k: An electronic structure and molecular dynamics software package - quickstep: Efficient and accurate electronic structure calculations," J. Chem. Phys. **152**, 194103 (2020).

[28] G. Bergerhoff, R. Hundt, R. Sievers, and I. D. Brown, "The inorganic crystal structure data base," J. Chem. Inf. Comput. Sci. **23**, 66–69 (1983).

[29] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, "Influence of the exchange screening parameter on the performance of screened hybrid functionals," J. Chem. Phys. **125**, 224106 (2006).

[30] J. L. Lyons and C. G. Van de Walle, "Computationally predicted energies and properties of defects in GaN," npj Comput. Mater. **3**, 1–10 (2017).

[31] Y. K. Frodason, J. B. Varley, K. M. H. Johansen, L. Vines, and C. G. Van de Walle, "Migration of Ga vacancies and interstitials in β-Ga$_2$O$_3$," Phys. Rev. B **107**, 024109 (2023).

[32] J.-X. Shen, H. H. Li, A. Rutt, M. K. Horton, and K. A. Persson, "Topological graph-based analysis of solid-state ion migration," npj Comput. Mater. **9**, 1–5 (2023).

[33] M. E. Ingebrigtsen, A. Yu. Kuznetsov, B. G. Svensson, G. Alfieri, A. Mihaila, U. Badstübner, A. Perron, L. Vines, and J. B. Varley, "Impact of proton irradiation on conductivity and deep level defects in β-Ga$_2$O$_3$," APL Mater. **7**, 022510 (2019).

[34] J. B. Varley, H. Peelaers, A. Janotti, and C. G. Van de Walle, "Hydrogenated cation vacancies in semiconducting oxides," J. Phys.: Condens. Matter **23**, 334212 (2011).

[35] A. Kyrtsos, M. Matsubara, and E. Bellotti, "Migration mechanisms and diffusion barriers of vacancies in Ga$_2$O$_3$," Phys. Rev. B **95**, 245202 (2017).

20 April 2024 06:39:58

[36]B. Liu, V. R. Cooper, H. Xu, H. Xiao, Y. Zhang, and W. J. Weber, "Composition dependent intrinsic defect structures in SrTiO₃," Phys. Chem. Chem. Phys. **16**, 15590–15596 (2014).

[37]J. N. Baker, P. C. Bowes, D. M. Long, A. Moballegh, J. S. Harris, E. C. Dickey, and D. L. Irving, "Defect mechanisms of coloration in Fe-doped SrTiO₃ from first principles," Appl. Phys. Lett. **110**, 122903 (2017).

[38]A. Janotti, J. B. Varley, M. Choi, and C. G. Van de Walle, "Vacancies and small polarons in SrTiO₃," Phys. Rev. B **90**, 085202 (2014).

[39]J. B. Varley, A. Janotti, and C. G. Van de Walle, "Hydrogenated vacancies and hidden hydrogen in SrTiO₃," Phys. Rev. B **89**, 075202 (2014).

[40]M. Choi, F. Oba, and I. Tanaka, "Role of Ti antisitelike defects in SrTiO₃," Phys. Rev. Lett. **103**, 185502 (2009).

[41]A. Alkauskas, P. Broqvist, and A. Pasquarello, "Defect levels through hybrid density functionals: Insights and applications," Phys. Status Solidi B **248**, 775–789 (2011).

[42]S. Muy *et al.*, "AiiDA-defects: An automated and fully reproducible workflow for the complete characterization of defect chemistry in functional materials," Electron. Struct. **5**(2), 024009 (2023).

[43]J.-X. Shen and Joel Varley, "pymatgen-analysis-defects: A Python package for analyzing point defects in crystalline materials," J. Open Source Software **9**(93), 5941 (2024).

20 April 2024 06:39:58