Finite-size corrections to defect energetics along one-dimensional configuration coordinate

Yu Kumagai [*]

Institute for Materials Research, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan

Recently, effective one-dimensional configuration coordinate diagrams have been utilized to calculate the line shapes of luminescence spectra and nonradiative carrier capture coefficients via point defects. Their calculations necessitate accurate total energies as a function of configuration coordinates. Although supercells under periodic boundary conditions are commonly employed, the spurious cell size effects have not been previously considered. In this study, we have proposed a correction energy formalism and verified its effectiveness by applying it to a nitrogen vacancy in GaN.

Introduction. Point defects are ubiquitous in solids, and they play important roles in many fundamental physics (e.g., superconductors [1] and topological materials [2]) and practical applications (e.g., solar cell absorbers [3] and light-emitting diodes [4]). To better understand and even predict the point defect properties, first-principles calculations have been routinely used over the past few decades [5]. Recently, the line shapes of luminescence spectra [6,7] and carrier capture coefficients [8–11] via point defects have been calculated from first principles. The key approximation in these calculations is to simplify the defect-related phonon modes to one effective mode [8] that is parallel to the distortion of the defect geometry when gaining or losing an electron. Total energies as a function of such a one-dimensional (1D) configuration are schematically described with a configuration coordinate (CC) diagram as shown in Fig. 1.

To accurately determine the formation energies of charged defects using supercells under periodic boundary conditions (PBCs), it is necessary to account for finite cell size errors that result from spurious electrostatic interactions. Formalisms have been developed for making such corrections, and the Freysoldt-Neugebauer–Van de Walle (FNV) method [12] is currently recognized as the most advanced approach. We have expanded the FNV method to accommodate anisotropic materials with atomic relaxation (eFNV) [13]. Optical transition levels via point defects, also known as vertical transition (VT) energies, can be computed assuming that the atomic configuration remains unchanged during the electron transitions, in accordance with the Franck-Condon principle [14,15]. Their calculations require a distinct correction methodology, for which we have proposed a method termed the GKFO method [16]. Falletta *et al.* have also reported a comparable correction energy from a different viewpoint [17].

So far, when calculating the CC diagram, all the energies along the CC were constantly shifted by the correction energies obtained at the ground state using the FNV method.

Unfortunately, this approach is incorrect. To understand why, let us begin by defining some notation.

The atomic configuration R from the relaxed structure at charge q to the relaxed structure at charge $q + \Delta q$ ($\Delta q = \pm 1$) is denoted as $R^{q \to q + \Delta q}(\alpha)$, where the structure is linearly interpolated or extrapolated in a ratio of α. When $\alpha = 1$, the atomic configuration corresponds to the relaxed structure with charge $q + \Delta q$. The defect state at charge q and configuration $R^{q \to q + \Delta q}(\alpha)$ is denoted as $(q, R^{q \to q + \Delta q}(\alpha))$, and its energy as $E(q, R^{q \to q + \Delta q}(\alpha))$. We also define its relative energy with respect to $E(q, R^{q \to q + \Delta q}(0))$ as $\Delta E(q, R^{q \to q + \Delta q}(\alpha))$ (see Fig. 1).

In the case of $q = 0$ and $\Delta q = +1$, $E(0, R^{0 \to +1}(0))$ does not require any electrostatic correction. However, when, e.g., $\alpha = 1$, the neighboring ions move slightly as if $+1$ charge is located at the defect site. Such ion distortion occurs linearly with respect to α and is effectively regarded as a polarization charge [17], which interacts with its images under PBCs and requires energy corrections depending on α. Our objective is to develop a formalism for correcting these energies and to verify its effectiveness.

To obtain the correction energy for $\Delta E(q, R^{q \to q + \Delta q}(\alpha))$ (ΔE_{cor}), we introduce a virtual state $(q + \alpha \Delta q, R^{q \to q + \Delta q}(\alpha))$, in which the ionic dipoles are exactly balanced with the defect charge $q + \alpha \Delta q$. While computing such a state explicitly is computationally demanding, it is not necessary in the final formulation. ΔE_{cor} is the sum of the (i) eFNV and (ii) GKFO correction energies, as illustrated in Fig. 1. Initially, we consider the cubic system and later generalize it.

U_ε^q is the point-charge (PC) correction energy for charge q screened by a dielectric constant ε, and C_ε^q is an alignment constant chosen such that the short-range potential decays to zero far from the defect after removing the long-range potential caused by charge q and screened by ε (see Refs. [12,13]). As derived by Komsa *et al.* [18], the alignment term can be generally written as

$$C_\varepsilon^q = \frac{2\pi \int d\mathbf{r}^3 \Delta \rho^q r^2}{\varepsilon L^3}, \tag{1}$$

*yukumagai@tohoku.ac.jp

where L is the side length of the supercell, and $\Delta\rho^q$ is the defect charge distribution modified by adding charge q.

The correction energies for the relaxed structures with charge q and $q + \alpha\Delta q$ are then expressed in the eFNV approach as $U_{\varepsilon_0}^q - qC_{\varepsilon_0}^q$ and $U_{\varepsilon_0}^{q+\alpha\Delta q} - (q + \alpha\Delta q)C_{\varepsilon_0}^{q+\alpha\Delta q}$, respectively. Here, ε_0 is the static dielectric constant, which is the sum of the ion-clamped (ε_∞) and ionic (ε_{ion}) dielectric constants. The correction energy for route (i) is then described as

$$\Delta E_{\text{cor}}^{(i)} = U_{\varepsilon_0}^{q+\alpha\Delta q} - U_{\varepsilon_0}^q - \left[(q + \alpha\Delta q)C_{\varepsilon_0}^{q+\alpha\Delta q} - qC_{\varepsilon_0}^q\right]. \tag{2}$$

Route (ii) corresponds to the VT, where the charge transitions from $q + \alpha\Delta q$ to q, and it is corrected with the GKFO method as follows:

$$\begin{aligned}
\Delta E_{\text{cor}}^{(ii)} = &- \frac{2\alpha\Delta q}{q + \alpha\Delta q} U_{\varepsilon_0}^{q+\alpha\Delta q} + U_{\varepsilon_\infty}^{-\alpha\Delta q} \\
&- \left(- \alpha\Delta q C_{\varepsilon_\infty}^{-\alpha\Delta q} - \alpha\Delta q C_{\varepsilon_0}^{q+\alpha\Delta q} \right. \\
&\left. + \frac{\varepsilon_\infty}{\varepsilon_0}(q + \alpha\Delta q)C_{\varepsilon_\infty}^{-\alpha\Delta q}\right).
\end{aligned} \tag{3}$$

Here, $C_{\varepsilon_\infty}^{-\alpha\Delta q}$ is an alignment constant caused by introducing additional charge $-\alpha\Delta q$ without altering the atomic configuration $R^{q\to q+\Delta q}(\alpha)$ [16].

The sum of the PC correction terms can be expressed as

$$\Delta E_{\text{cor}}^{\text{PC}} = U_{\varepsilon_\infty}^{-\alpha\Delta q} - U_{\varepsilon_0}^{-\alpha\Delta q} \tag{4}$$

(see Supplemental Material [19] for the detailed derivation). Because $\Delta q = \pm 1$ and the sign of Δq is irrelevant for $U^{-\alpha\Delta q}$, $\Delta E_{\text{cor}}^{\text{PC}}$ depends on neither initial charge q nor Δq but on α. Note, however, that ΔE_{cor} represents the additional correction to the *relative* energy with respect to the relaxed structure, and the total correction energy is the sum of ΔE_{cor} and the FNV correction energy to the relaxed structure, where the PC term depends on q^2.

The summation of the alignment terms is written as

$$\begin{aligned}
\Delta E_{\text{cor}}^{\text{align}} = &- (q + \alpha\Delta q)C_{\varepsilon_0}^{q+\alpha\Delta q} + qC_{\varepsilon_0}^q \\
&+ \alpha\Delta q C_{\varepsilon_\infty}^{-\alpha\Delta q} + \alpha\Delta q C_{\varepsilon_0}^{q+\alpha\Delta q} \\
&- \frac{\varepsilon_\infty}{\varepsilon_0}(q + \alpha\Delta q)C_{\varepsilon_\infty}^{-\alpha\Delta q}.
\end{aligned} \tag{5}$$

When we define ε_{eff} as $1/\varepsilon_{\text{eff}} = 1/\varepsilon_\infty - 1/\varepsilon_0$, $\Delta E_{\text{cor}}^{\text{align}}$ is found to be $\alpha\Delta q C_{\varepsilon_{\text{eff}}}^{-\alpha\Delta q}$ (see Supplemental Material [19] for details of the derivation). Thus, ΔE_{cor} is expressed as

$$\Delta E_{\text{cor}} = U_{\varepsilon_{\text{eff}}}^{-\alpha\Delta q} - (-\alpha\Delta q)C_{\varepsilon_{\text{eff}}}^{-\alpha\Delta q}. \tag{6}$$

The obtained ΔE_{cor} is simply estimated by considering the effective charge $-\alpha\Delta q$ screened by an effective dielectric constant ε_{eff}. This can be easily extended to the anisotropic systems by replacing ε_{eff} with the tensor form, $\bar{\varepsilon}_{\text{eff}}$. Furthermore, it is applicable to the anharmonic potential energy surface [9] because the ionic displacement distance for the screening depends linearly on α.

To verify the effectiveness of the correction formalism in Eq. (6), we have calculated the total energies of the nitrogen vacancy (V_{N}) in GaN. To demonstrate that the initial charge
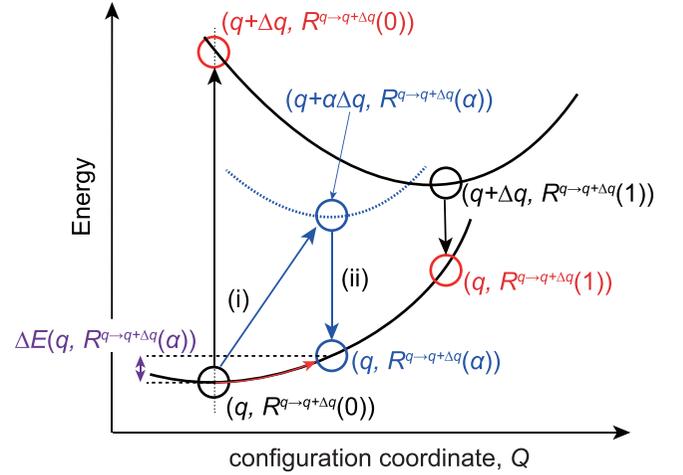
FIG. 1. Schematics of the configuration coordinate diagram between charge q and $q + \Delta q$, where $\Delta q = \pm 1$. The black circles depict energies of structures relaxed at each charge, whereas the red circles represent those relaxed at different charges. The blue circles denote the energy in virtual state $(q + \alpha\Delta q, R^{q\to q+\Delta q}(\alpha))$ (see the text for details). The correction energy for $\Delta E(q, R^{q\to q+\Delta q}(\alpha))$ is equal to the sum of correction energies for routes (i) and (ii).

state does not affect ΔE_{cor}, we calculated the energies at $q = +2$ along the CC to the structure at $q = +3$. First, we confirm that the conventional eFNV method works well for calculating the formation energy of V_{N}^{+2} [$E(V_{\text{N}}^{+2})$] in GaN, especially when using the 300-atom supercell, as shown in Fig. 2(a).

In Figs. 2(b) and 2(c), we present the atomic site potentials caused by V_{N}^{+2} at two different configuration coordinates, namely $\alpha = 0.5$ and 1.5, in the 576-atom supercell. Despite having identical charge $q = +2$, a significant difference in the potential variation is observed between the two coordinates. The atomic site potentials calculated from a charge $-\alpha\Delta q$, where $\Delta q = +1$, at the defect site screened by $\bar{\varepsilon}_{\text{eff}}$ are also displayed. These potentials adequately reproduce tendencies of the atomic site potentials in the region far from the defect, leading to an estimation of $C_{\varepsilon_{\text{eff}}}^{-\alpha\Delta q}$.

Figure 2(d) presents $\Delta E(2, R^{2\to 3}(1.5))$ of V_{N}^{+2} as a function of the number of atoms in the pristine supercell (N_{atoms}). When applying the constant energy correction irrelevant to the CC that corresponds to "Uncorrected" in this Letter, there are noticeable finite-cell size errors. These errors are effectively eliminated when ΔE_{cor} in Eq. (6) are applied, thereby validating our formalism. As shown in Fig. 2(d), the alignment terms do not exert a significant effect, meaning that the ionic dipoles that screen the additional defect charge do not appreciably alter the defect charge.

Figure 2(e) depicts $\Delta E(2, R^{2\to 3}(\alpha))$ with and without corrections, as a function of the displacement fraction α. As Eq. (6) suggests, the finite-size errors are predominantly dependent on α^2. When computing the nonradiative carrier capture coefficients, the position at which the potential energy surfaces of two states intersect has a significant effect on the capture rate. Such intersections can occur at values of α greater than 2, as exemplified by the carbon impurity at the nitrogen site in GaN [8]. Hence, it is likely that
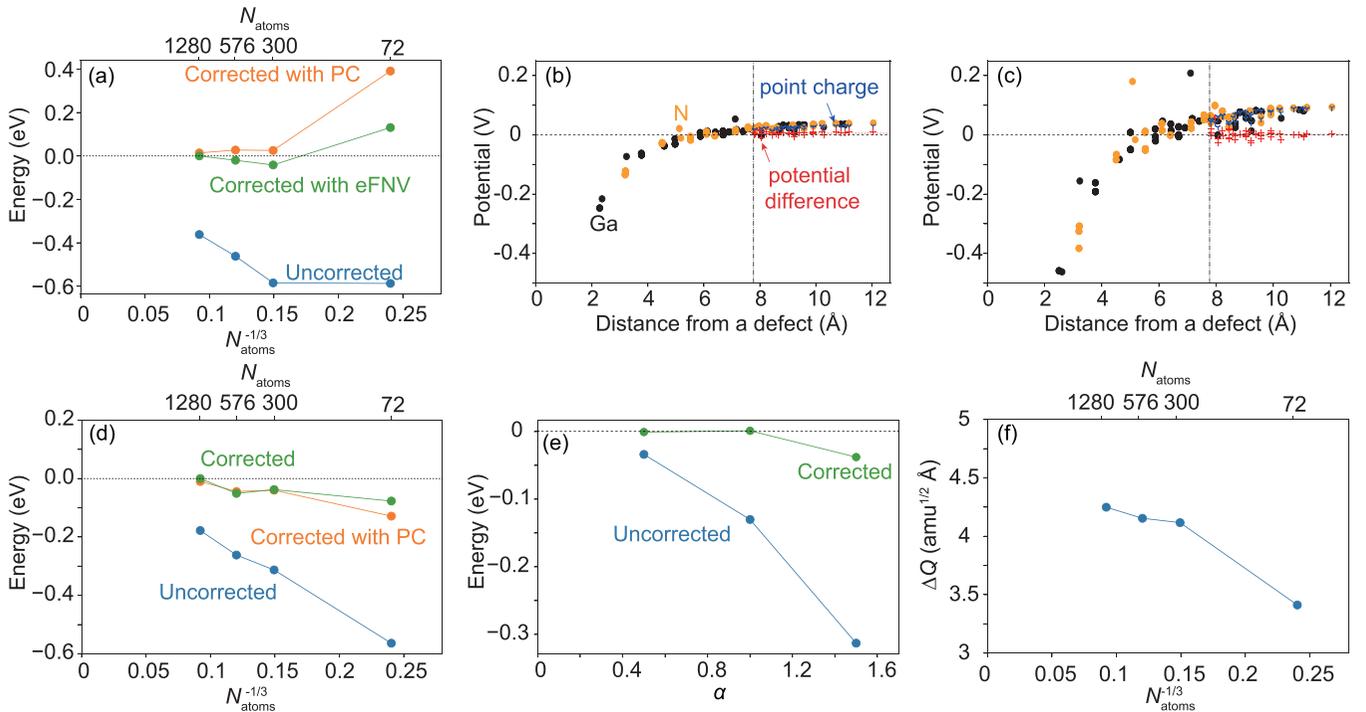
FIG. 2. (a) Relative formation energies of V_N^{+2} in GaN without any corrections, with the PC corrections, and with the eFNV corrections as a function of the supercell size. The zero-energy value is set to the energy corrected with the eFNV method at the largest supercell. (b),(c) Atomic site potentials caused by V_N^{+2} at (b) $(2, R^{2\to3}(0.5))$ and (c) $(2, R^{2\to3}(1.5))$ in the 576-atom supercell that are calculated by subtracting the potentials in the supercell containing V_N^{+2} at $(2, R^{2\to3}(0))$. Blue marks are the electrostatic potentials caused by periodic point charges $-\alpha\Delta q$ and background charge screened by an effective dielectric tensor, while red crosses are their differences from the atomic site potentials. The vertical dashed lines indicate the radius of the sphere in contact with the supercell. The averaged potential difference far from the defect $(C_{\varepsilon_{\mathrm{eff}}}^{-\alpha\Delta q})$ is evaluated outside of the sphere (see Ref. [13] for details). (d) The supercell size dependence of $\Delta E(2, R^{2\to3}(1.5))$ without any corrections, with PC corrections [Eq. (4)], and with corrections in Eq. (6). (e) $\Delta E(2, R^{2\to3}(\alpha))$ as a function of α without and with corrections in Eq. (6) calculated using the 300-atom supercell. Since the contributions of the alignment terms are only a few meV, the PC correction results are nearly identical to the fully corrected results. (f) ΔQ of V_N in GaN from $+2$ charge to $+3$ charge (see the text for details). In (d) and (e), the zero-energy values are set to the energies corrected with Eq. (6) at the largest supercells.

the intersection coordinates and energies are significantly modified by applying ΔE_{cor}. It is also worth noting that a miscalculation of 0.1 eV in the energies along the CC can result in a tenfold error in the nonradiative carrier capture coefficients [8].

As shown in Figs. 2(a) and 2(d), $E(V_N^{+2})$ and $\Delta E(2, R^{2\to3}(1.5))$ are effectively corrected even using the 72-atom supercell. However, the difference of the configuration coordinate Q (see the Method section) between charge q and $q + \Delta q$ (ΔQ) is not accurately estimated using the 72-atom supercell as shown in Fig. 2(f). This might also be related to the fact that the 72-atom supercell calculation of $E(V_N^{+2})$ shows a different tendency in Fig. 2(a). Because errors in ΔQ have a significant impact on the calculations of the line shapes of luminescence spectra and carrier capture coefficients, we generally recommend using larger supercells, particularly to improve ΔQ.

The effective dielectric constant $\varepsilon_{\mathrm{eff}}$ quantifies the polarization resulting from ion displacement. It is related to the long-range electron-phonon coupling, and it is commonly used when discussing self-trapped polarons [20]. In general, decreasing $\varepsilon_{\mathrm{eff}}$ leads to an increase in E_{cor}. Because $\varepsilon_{\mathrm{eff}} = \varepsilon_\infty + \varepsilon_\infty^2/\varepsilon_{\mathrm{ion}}$, $\varepsilon_{\mathrm{eff}}$ decreases as ε_∞ decreases and $\varepsilon_{\mathrm{ion}}$ increases. To assess the potential magnitude of the error in

realistic materials resulting from neglecting ΔE_{cor} in Eq. (6), we calculated $\varepsilon_{\mathrm{eff}}$ for 931 oxide materials for which we calculated the oxygen vacancies in our previous work [21], using the spherical averages of ε_∞ and $\varepsilon_{\mathrm{ion}}$. The smallest $\varepsilon_{\mathrm{eff}}$, which is 2.5, is found in $CsLi_5(BO_3)_2$, where the averaged ε_∞ and $\varepsilon_{\mathrm{ion}}$ are 2.1 and 10.8, respectively. Then, ΔE_{cor} for the oxygen vacancy at state $(1, R^{1\to2}(\alpha = 2))$ is estimated to be 1.2 eV even using the very large 448-atom supercell.

We can estimate how much the energy correction modifies the effective phonon frequency, which is given by the equation $\Omega^2 = \frac{\partial^2 E}{\partial Q^2}$ in the harmonic approximation [6,8]. Replacing Ω with $\Omega + \Omega_{\mathrm{cor}}$, where Ω_{cor} is the corrected frequency, and E with $E + \Delta E_{\mathrm{cor}}$, and assuming that $\Omega_{\mathrm{cor}} \ll \Omega$ and the PC term is dominant in Eq. (6), we obtain

$$\Omega_{\mathrm{cor}} = \frac{U_{\varepsilon_{\mathrm{eff}}}^{\Delta q}}{\Omega(\Delta Q)^2} \tag{7}$$

(see Supplemental Material [19] for details of the derivation). For the case of $(2, R^{2\to3}(\alpha))$ for V_N, we find $\hbar\Omega_{\mathrm{cor}} = 3.28$ meV when using the 72-atom supercell, which is 14.5% of the uncorrected $\hbar\Omega$. Note that, because $\hbar\Omega_{\mathrm{cor}}$ is approximately proportional to $U_{\varepsilon_{\mathrm{eff}}}^{\Delta q}$, its convergence as a function of N_{atom} is slow if ΔE_{cor} is not applied.

To conclude, we derived the correction formalism for the defect energies along the 1D CC, which has not been considered so far. The correction energy is described with the effective charge $-\alpha\Delta q$ and effective dielectric constant ε_{eff}. Its effectiveness has been verified with a nitrogen vacancy in GaN. Since our correction method is automatically applied with negligible computational cost compared to first-principles calculations, we believe that it will be routinely used in the future.

Method. All the calculations were performed using the projector augmented-wave (PAW) method [22,23] implemented in VASP [24]. We adopted the HSE06 hybrid functional [25]. A fraction of Fock exchange parameter was set to 0.31 to reproduce the experimental band gap [8]. See Supplemental Material [19] for the computational details. The harmonic phonon frequencies were calculated using NONRAD [26]. All the VASP input settings were generated with the VISE code (version 0.7.0) [27], while the processing related to defects was done with PYDEFECT [28]. The difference of the configuration coordinate ΔQ is defined as $\sqrt{\sum_\alpha M_\alpha \cdot \Delta\boldsymbol{R}_\alpha^2}$, where M_α and $\Delta\boldsymbol{R}_\alpha$ are the atomic mass and the displacement vector from the equilibrium position of atom α, respectively [6,8]. The ion-clamped dielectric constants (ε_∞) for 931 oxides were obtained from the long-range wavelength limit in the real part of dielectric functions calculated using the dielectric dependent hybrid functional [29] (see Ref. [30] for details). The ionic contributions (ε_{ion}) were calculated using the density functional perturbation theory [31] (see Ref. [21] for details).

[1] J. G. Bednorz and K. A. Müller, Possible high Tc superconductivity in the Ba–La–Cu–O system, Z. Phys. B **64**, 189 (1986).

[2] Y. Ando, Topological insulator materials, J. Phys. Soc. Jpn. **82**, 102001 (2013).

[3] P. K. Nayak, S. Mahesh, H. J. Snaith, and D. Cahen, Photovoltaic solar cell technologies: Analysing the state of the art, Nat. Rev. Mater. **4**, 269 (2019).

[4] S. Nakamura, N. Iwasa, M. S. M. Senoh, and T. M. T. Mukai, Hole compensation mechanism of P-type GaN films, Jpn. J. Appl. Phys. **31**, 1258 (1992).

[5] 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 (2014).

[6] A. Alkauskas, J. L. Lyons, D. Steiauf, and C. G. Van de Walle, First-Principles Calculations of Luminescence Spectrum Line Shapes for Defects in Semiconductors: The Example of GaN and ZnO, Phys. Rev. Lett. **109**, 267401 (2012).

[7] Y. K. Frodason, K. M. Johansen, T. S. Bjørheim, B. G. Svensson, and A. Alkauskas, Zn vacancy as a polaronic hole trap in ZnO, Phys. Rev. B **95**, 094105 (2017).

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

[9] S. Kim, S. N. Hood, and A. Walsh, Anharmonic lattice relaxation during nonradiative carrier capture, Phys. Rev. B **100**, 041202(R) (2019).

[10] X. Zhang, M. E. Turiansky, and C. G. Van de Walle, Correctly assessing defect tolerance in halide perovskites, J. Phys. Chem. C **124**, 6022 (2020).

[11] S. R. Kavanagh, D. O. Scanlon, A. Walsh, and C. Freysoldt, Impact of metastable defect structures on carrier recombination in solar cells, Faraday Discuss. **239**, 339 (2022).

[12] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, Fully *Ab Initio* Finite-Size Corrections for Charged-Defect Supercell Calculations, Phys. Rev. Lett. **102**, 016402 (2009).

[13] Y. Kumagai and F. Oba, Electrostatics-based finite-size corrections for first-principles point defect calculations, Phys. Rev. B **89**, 195205 (2014).

[14] J. Franck and E. G. Dymond, Elementary processes of photochemical reactions, Trans. Faraday Soc. **21**, 536 (1926).

[15] E. Condon, A theory of intensity distribution in band systems, Phys. Rev. **28**, 1182 (1926).

[16] T. Gake, Y. Kumagai, C. Freysoldt, and F. Oba, Finite-size corrections for defect-involving vertical transitions in supercell calculations, Phys. Rev. B **101**, 020102(R) (2020).

[17] S. Falletta, J. Wiktor, and A. Pasquarello, Finite-size corrections of defect energy levels involving ionic polarization, Phys. Rev. B **102**, 041115(R) (2020).

[18] H.-P. Komsa, T. T. Rantala, and A. Pasquarello, Finite-size supercell correction schemes for charged defect calculations, Phys. Rev. B **86**, 045112 (2012).

[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.107.L220101 for the detailed derivation of the equations shown in the main text, and detailed computational settings.

[20] D. Emin, *Polarons* (Cambridge University Press, Cambridge, 2013).

[21] 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).

[22] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).

[23] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B **59**, 1758 (1999).

[24] 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 (1996).

[25] 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).

[26] M. E. Turiansky, A. Alkauskas, M. Engel, G. Kresse, D. Wickramaratne, J.-X. Shen, C. E. Dreyer, and C. G. Van de Walle, Nonrad: Computing nonradiative capture coefficients from first principles, Comput. Phys. Commun. **267**, 108056 (2021).

[27] Y. Kumagai, vise (2021), https://github.com/kumagai-group/vise.

[28] Y. Kumagai, pydefect (2021), https://github.com/kumagai-group/pydefect.

[29] J. H. Skone, M. Govoni, and G. Galli, Self-consistent hybrid functional for condensed systems, Phys. Rev. B **89**, 195112 (2014).

[30] Y. Kumagai, Computational Screening of p-Type Transparent Conducting Oxides Using the Optical Absorption Spectra and Oxygen-Vacancy Formation Energies, Phys. Rev. Appl. **19**, 034063 (2023).

[31] X. Gonze and C. Lee, Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory, Phys. Rev. B **55**, 10355 (1997).