Convergence of density and hybrid functional defect calculations for compound semiconductors

Haowei Peng,[1] David O. Scanlon,[2,3] Vladan Stevanovic,[1,4] Julien Vidal,[5] Graeme W. Watson,[6] and Stephan Lany[1]
[1]*National Renewable Energy Laboratory, Golden, Colorado 80401, USA*
[2]*University College London, Kathleen Lonsdale Materials Chemistry, Department of Chemistry, 20 Gordon Street, London WC1H 0AJ, United Kingdom*
[3]*Diamond Light Source Ltd., Diamond House, Harwell Science and Innovation Campus, Didcot, Oxfordshire OX11 0DE, United Kingdom*
[4]*Colorado School of Mines, Department of Physics, Golden, Colorado 80401, USA*
[5]*Institute for Research and Development of Photovoltaic Energy (IRDEP), UMR 7174 CNRS/EDF R&D/Chimie ParisTech, 6 quai Watier, 78401 Chatou, France*
[6]*School of Chemistry and CRANN, Trinity College Dublin, Dublin 2, Ireland*
(Received 31 May 2013; revised manuscript received 29 July 2013; published 3 September 2013)

Recent revisions of defect formation energy calculations based on bandgap corrected hybrid functionals have raised concerns about the validity of earlier results based on standard density functionals and about the reliability of the theoretical prediction of electrical properties in semiconductor materials in general. We show here that a close agreement between the two types of functionals can be achieved by determining appropriate values for the electronic and atomic reference energies, thereby mitigating uncertainties associated with the choice of the underlying functional.

DOI: 10.1103/PhysRevB.88.115201 PACS number(s): 71.15.−m, 71.20.Nr, 71.55.−i

I. INTRODUCTION

The theoretical prediction of the electrical properties of semiconductors relies on the ability to calculate accurate defect formation energies ΔH. The methodology of supercell calculations for ΔH is a topic that continues to receive great interest.[1–5] A major issue that has plagued such calculations for a long time is the "bandgap problem" of the local density and generalized gradient approximations (LDA and GGA) within density functional theory (DFT). Particularly for charged defects, such as electrically active dopants, for which ΔH depends on the Fermi level, an unambiguous prediction of the formation energy requires the correct bandgap energy, and a number of correction schemes have been discussed and applied to map the formation energies and transition levels within the underestimated DFT gap onto the full gap.[1,6–8] A major paradigmatic change in the field took place when the implementation of Fock exchange into plane wave DFT codes[9,10] allowed bandgap corrected hybrid functional[11,12] calculations of supercells, albeit at the cost of a considerably increased computational overhead that sometimes invites compromises on convergence parameters. A considerable number of works have since been devoted to revisit previously studied cases, often finding quantitatively different results or even a qualitatively changed physical picture.[13–18] Such revisions have raised concerns about the validity of previous density functional results and about the accuracy of ΔH calculations in general. Considering a set of six semiconductor materials (Cu_3N, Cu_2O, Zn_3N_2, ZnO, AlN, and Al_2O_3) and a total of 24 defects and dopants in their fully ionized charge states, we show here that the difference between standard density and hybrid functionals reflects systematically the dependence of ΔH on the reference energies for the electronic and atomic chemical potentials. The convergence of the ΔH predicted by the two types of functionals is achieved by using appropriate values for these reference energies from the one-particle Green function G and the screened Coulomb interaction W (GW) quasiparticle energy calculations[19] and from the fitted elemental reference energies (FERE) approach,[20,21] respectively. By using the GW band edge shifts to predict the absolute formation energies of fully ionized dopants and defects, this paper complements previous work, where such shifts were applied to band offsets[22,23] and the alignment of deep-level transition energies.[24]

While hybrid functionals allow one to tune the band gap via the mixing parameter α for the Fock exchange and the range separation parameter in the case of the Heyd, Scuseria, Ernzerhof (HSE) functional,[12] there are several potential sources of uncertainty associated with the choice of these parameters: (i) The parameters needed to reproduce the experimental band gap vary significantly, e.g. from $\alpha = 0.20$ in TiO_2 (Ref. 17) to $\alpha = 0.38$ for ZnO (Ref. 13) (both with a HSE range separation parameter of 0.2 $Å^{-1}$), introducing some ambiguity for defect calculations, e.g. if one were to consider a Ti defect in ZnO or vice versa. (ii) The parameters chosen for the semiconductor compound are not well justified for the elemental phases (e.g. the metallic phases of Ti or Zn and the O_2 molecule), which could affect the thermodynamic limits that define the range of defect formation energies as a function of the chemical potentials. (iii) The parameters that correct the band gap do not necessarily correct at the same time also other features, e.g. the correct d-band position in ZnO,[25] the linearity of the total energy (Koopmans behavior) needed to describe correctly localized polaronic electron or hole states,[26,27] or the individual absolute band edge energies, which are needed, e.g. for band offset calculations.[22,23] Thus, it is desirable to devise approaches in which the defect formation energy is not strongly dependent on the hybrid functional parameters.

II. ELECTRONIC AND ATOMIC REFERENCE ENERGIES

The formation energy ΔH of a defect D in the charge state q is usually expressed as a sum of three terms

$$\Delta H(E_F,\{\mu_a\}) = \left[E_{\rm tot}^{D,q} - E_{\rm tot}^{H}\right] + q\left(E_{\rm VBM}^{\rm ref} + \Delta E_F\right) + \sum_a n_a\left(\mu_a^{\rm ref} + \Delta\mu_a\right), \quad (1)$$

i.e. it is a function of the Fermi energy $E_F = E^{\rm ref}_{\rm VBM} + \Delta E_F$ and the chemical potentials $\mu_a = \mu^{\rm ref}_a + \Delta\mu_a$, and n_a is the number of atoms of type a removed from ($n > 0$) or added to ($n < 0$) the crystal to create the defect. Taken by itself, the first term $[E^{D,q}_{\rm tot} - E^{H}_{\rm tot}]$, i.e. the total energy difference between the host matrix with and without the defect, is not particularly useful, since the total energies correspond to different numbers of atoms and electrons (for $q \neq 0$). A useful form of the defect formation energy is obtained only once the references for the Fermi energy and the chemical potentials are defined by physically meaningful limits. Conventionally, these limits are taken as the energy $E^{\rm ref}_{\rm VBM}$ of the valence band maximum (VBM) in the second term of Eq. (1), and as the (zero-temperature) chemical potential $\mu^{\rm ref}_a$ of the elemental phases of the atomic species a in the third term. For purpose of this paper, we define the reference defect formation enthalpy

$$\Delta H_{\rm ref} = \left[E^{D,q}_{\rm tot} - E^{H}_{\rm tot}\right] + q E^{\rm ref}_{\rm VBM} + \sum_a n_a \mu^{\rm ref}_a, \qquad (2)$$

which does not depend on the variable ranges ΔE_F and $\Delta\mu_a$ of the Fermi level and chemical potentials.

The notorious bandgap problem[28] in DFT leads to an ambiguity of ΔH for charged defects through the electronic reference energy $E^{\rm ref}_{\rm VBM}$. By its underlying principles,[29] DFT is constructed to give the correct charge density (and, hence, the correct electrostatic potential), as well as the correct total energy. However, the single-particle (Kohn-Sham) energies do not have the meaning of quasiparticle (electron-removal or electron-addition) energies, and LDA or GGA calculations underestimate the band gap, often by more than 50%. Note that an exact DFT functional would predict the correct band gap from the total energy difference $E_g = E(N+1) + E(N-1) - 2E(N)$, irrespective of the magnitude of the single particle gap.[28] In LDA or GGA, however, the total-energy and single-particle-energy gaps must coincide due to the absence of a discontinuity in these exchange-correlation potentials. While the closed-shell energy $E(N)$ is not directly affected by the band gap problem, the open-shell energies $E(N-1)$ and $E(N+1)$ are, which is recently being discussed in the context of the "delocalization error".[30] Thus, DFT approximations should describe rather accurately also the formation energies of fully ionized (closed-shell) defects, relative to the average electrostatic potential V_{av}, which defines the energy zero in periodic plane-wave methods.[31] As illustrated graphically in Fig. 1, an accurate prediction of the charged-defect formation energy should therefore be accessible by correcting the charged-defect formation energy using the GW quasiparticle energy shift $\delta E_{\rm VBM} = e^{GW}_{\rm VBM} - e^{\rm DFT}_{\rm VBM}$,

$$\Delta H^{GW}_{\rm ref} = \Delta H^{\rm DFT}_{\rm ref} + q\delta E_{\rm VBM}. \qquad (3)$$

The need to attribute the bandgap correction to the VBM and the conduction band minimum (CBM) in a certain proportion was recognized before,[1,6–8] but there has been so far little clarity how to appropriately determine the individual band edge shifts for charged-defect formation energies. In this paper, ΔH of dopants and defects in their fully ionized charge state is anchored at the average potential V_{av} and the GW quasiparticle energies are used to determine the appropriate VBM reference energy. In the fully ionized state there are no occupied defect levels (donor states populated by electrons or acceptor states populated by holes), which could require a post-DFT treatment with a contribution to the formation energy. The present approach is fully compatible with the use of the same band edge shifts for band offsets[22,23] and the positioning of deep level transition energies[24] (proposed that these levels are indeed invariant relative to V_{av} as found for a variety of defects in Ref. 24). Conceptually also related is the work of Ref. 4, where the vacuum level was used as a reference for the band edge energies in the context of defect calculations. Noting that the charged-defect formation energies do not explicitly depend on the potential step at the semiconductor-vacuum interface, we agree that integration into a common picture is possible since the abovementioned DFT principles imply a consistent description of both defect energies and the potential step.

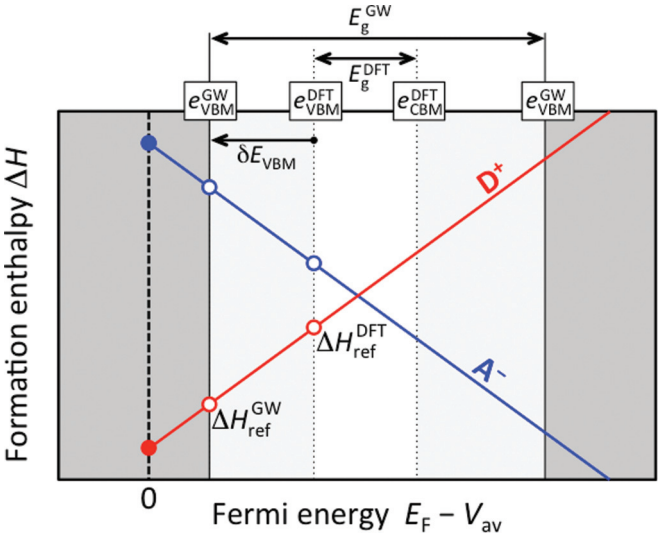

FIG. 1. (Color online) Within DFT, the formation energy ΔH for the fully ionized defects should be rather accurately described relative to the average electrostatic potential V_{av}, but the single-particle energies of the band edges ($e^{\rm DFT}_{\rm VBM}$ and $e^{\rm DFT}_{\rm CBM}$) are usually not. The quasiparticle energies $e^{GW}_{\rm VBM}$ and $e^{GW}_{\rm CBM}$ define the corrected range of $\Delta H(E_F)$. Schematically shown are ΔH for a positively charged (donorlike) defect D and a negatively charged (acceptorlike) defect A.

In a similar way as $E^{\rm ref}_{\rm VBM}$ defines the reference energy for the electron reservoir (Fermi level), the elemental energies $\mu^{\rm ref}_a$ define the reference energy for the atomic reservoir [cf. Eqs. (1)–(2)]. At the same time, the $\mu^{\rm ref}_a$ enter the calculation of the compound formation enthalpy

$$\Delta H^C_f = E^C_{\rm tot} - \sum_a m_a \mu^{\rm ref}_a, \qquad (4)$$

which is needed to determine the thermodynamic limits of the chemical potentials $\Delta\mu_a$ [cf. Eq. (1)]. Here, $E^C_{\rm tot}$ is the total energy of the compound consisting of m_a atoms of type a per formula unit. Standard LDA or GGA calculations have been found to predict ΔH^C_f with less than desirable accuracy.[20,21,32,33] The discrepancy of 0.2–0.3 eV/atom compared to experiment is not as dramatic as the LDA bandgap error, but sufficiently large to merit the development of approaches for correction.[20,21,32,33] The origin of these inaccuracies has been traced back to incomplete error cancellation between chemically dissimilar

TABLE I. The change of the band gap and of the VBM energy due to GW quasiparticle energy corrections relative the GGA($+U$) and HSE. All numbers in eV.

	GGA($+U$)			HSE		
	E_g (DFT)	E_g (GW)	$\delta E_{\rm VBM}$	E_g (DFT)	E_g (GW)	$\delta E_{\rm VBM}$
Cu_3N	0.76	1.04	−0.44	0.98	0.99	−0.26
Zn_3N_2	0.20	1.00	−0.43	0.89	1.16	−0.12
AlN	4.21	6.29	−1.29	5.59	6.50	−0.45
Cu_2O	0.75	2.07	−0.67	2.03	2.35	−0.40
ZnO	1.52	3.38	−1.09	2.51	3.65	−0.55
Al_2O_3	6.26	9.78	−1.92	8.11	9.95	−0.70

systems (e.g. compounds, metals, molecules).[20] The fitted elemental-phase reference energies $\mu_a^{\rm FERE}$ were introduced in Refs. 20 and 21 to determine atomic reference energies that are compatible with the total energies calculated for the compounds in a given functional. The shifts $\delta\mu_a = \mu_a^{\rm FERE} - \mu_a^{\rm DFT}$ of the FERE relative to the directly calculated elemental energies in DFT are determined by the least-square solution of the set of linear equations

$$\sum_a m_a \delta\mu_a = \Delta H_f^{\rm DFT} - \Delta H_f^{\rm ex}, \qquad (5)$$

for a set of compounds that can be formed from a set of elements $\{a\}$. Improved estimates for the exact formation enthalpies $\Delta H_f^{\rm ex}$, needed to solve Eq. (5) for $\delta\mu_a$, could be determined from appropriate post-DFT total energy methods, such as quantum Monte Carlo,[34] or the random-phase approximation.[35,36] In this paper, we use instead the available tabulated experimental data[37] for binary and ternary nitrides and oxides of Cu, Al, and Zn. We note that the resulting elemental reference energies $\mu_a^{\rm ref} = \mu_a^{\rm FERE}$ do not depend on the DFT energies of the elemental phases.[20] Therefore, the abovementioned ambiguity, concerning the use of parameters (α in hybrid functionals, U in GGA $+ U$) for materials other than those they have been adjusted for, is removed. The final result for the defect formation energies with correction of both the electronic and atomic reference energies reads

$$\Delta H_{\rm ref}^{GW,\rm FERE} = \Delta H_{\rm ref}^{\rm DFT} + q\delta E_{\rm VBM} + \sum_a n_a \delta\mu_a. \qquad (6)$$

III. GW QUASIPARTICLE ENERGY CALCULATIONS

The present calculations were performed using the projector augmented wave (PAW)[38] implementation of DFT and GW in the VASP code.[9,39,40] For the DFT calculations, we use the GGA($+U$)[41,42] ($U = 5$ and 6 eV for Cu-d and Zn-d, respectively) and the HSE[12] ($\alpha = 0.25$) functionals. The GW calculations follow the scheme applied in the previous work of Refs. 25, 27, and 43. The GGA $+ U$ or HSE wave functions are maintained to keep the same average potential V_{av} as reference for the single- (quasi-) particle energies. The tendency of overestimated d-band energies is compensated by an attractive on-site potential for Zn-d [$V_d = -1.5$ eV (Ref. 25)] and Cu-d [−2.4 eV (Ref. 43)]. Further details and the computational parameters are given in the Appendix. As shown in Table I, the GW band gaps obtained with GGA($+U$) and HSE wave functions differ by less than 0.3 eV for all cases.

IV. FITTED ELEMENTAL REFERENCE ENERGIES

Table II gives the shifts $\delta\mu_a$ of the elemental reference energies obtained from the least-squares fit of Eq. (5), and Table III shows the compound formation enthalpies ΔH_f before and after the FERE correction. The root-mean-square (RMS) deviation from the experimental data is 0.28 and 0.22 eV/atom in the direct GGA($+U$) and HSE calculations, respectively, indicating that the thermochemistry is only slightly described better in HSE than in GGA($+U$). With the $\mu_a^{\rm FERE}$ reference energies, the deviation reduced to 0.02 and 0.01 eV/atom, respectively, which suggests that either functional describes excellently the relative energies between the different oxide and nitride compounds. The absolute $\mu_a^{\rm FERE}$ energies (see Appendix) for GGA($+U$) agree within less than 0.1 eV with previous results[20,21] that were obtained by fitting of a larger set of elements and compounds. Table II further reveals that the FERE shifts in the two functionals are almost identical for Cu, N, and O, but differ significantly for Zn and Al.

V. DEFECT FORMATION ENERGIES

For the calculation of the formation energies of the intrinsic and extrinsic defects, we are focusing here only on the important case of fully ionized defects or dopants, which usually define the defect equilibrium and, therefore, determine the net doping and the carrier densities. These highest stable charge states also exhibit the strongest dependence on the band edge shift $\delta E_{\rm VBM}$ [cf. Eq. (6) and Fig. 1]. Analogous corrections for the finite-size effects[1–3,44] have been applied in the case of either functional. The formation energies of all other charge states follow from the charge transition levels inside the band gap. For shallow defects, the ionization energy depends largely on the effective mass and the dielectric constant and, therefore, is not directly affected by the DFT bandgap error. In this case, the transition level can be easily determined from the ionization energy relative to the respective band edge. When

TABLE II. The shifts $\delta\mu_a$ of the elemental reference energies $\delta\mu_a = \mu_a^{\rm FERE} - \mu_a^{\rm DFT}$, given in electronvolts, relative to the directly calculated elemental energies in GGA($+U$) and HSE.

	Cu	Zn	Al	N	O
GGA($+U$)	−0.02	−0.08	+0.66	−0.03	+0.34
HSE	±0.00	+0.24	+0.35	−0.04	+0.32

TABLE III. The calculated compound formation enthalpies (electronvolts per formula unit) in GGA($+U$) and HSE, compared to experiment.[37]

		GGA($+U$)		HSE	
	ΔH_f (experiment)	ΔH_f (DFT)	ΔH_f (FERE)	ΔH_f (DFT)	ΔH_f (FERE)
Cu_3N	+0.77	+0.80	+0.89	+0.80	+0.85
Zn_3N_2	−0.23	−0.48	−0.18	+0.39	−0.26
AlN	−3.30	−2.73	−3.37	−3.01	−3.32
CuO	−1.63	−1.37	−1.69	−1.36	−1.68
Cu_2O	−1.75	−1.49	−1.80	−1.47	−1.77
ZnO	−3.61	−3.38	−3.64	−3.04	−3.60
Al_2O_3	−17.37	−14.90	−17.26	−15.68	−17.32
Al_2CuO_4	−18.79	−16.10	−18.77	−16.87	−18.82
$CuAlO_2$	−9.71	−8.33	−9.66	−8.71	−9.68
Al_2ZnO_4	−21.40	−18.73	−21.35	−19.19	−21.39

the ionization energy is increased due to polaronic localization, post-DFT corrections beyond LDA/GGA or possibly even beyond hybrid functionals need to be addressed, e.g. via a generalized Koopmans condition.[26] Other deep defects, such as typical negative-U defects, have defect states that are not directly related to a band edge state. In this case, the transition levels should be referred to the average potential V_{av} (cf. Fig. 1), but, again, post-DFT treatment of these levels, e.g. via defect GW calculations[45,46] might be necessary to obtain accurate transition energies.

Figure 2 shows the differences of the defect formation energies between HSE and the GGA($+U$) predictions before and after taking into account the GW and FERE corrections for the electronic and atomic reference energies, respectively. While the uncorrected values exhibit large differences with a RMS of 1.53 eV, the GW corrections bring the results of the two functionals into much closer agreement with an RMS of 0.46 eV, and the FERE corrections afford a further reduction of the difference to 0.34 eV. Since only the FERE shifts of Zn and Al exhibit a significant difference between GGA($+U$) and HSE (cf. Table II), the most pronounced change due to the FERE is observed in Fig. 2 for the cation substitutional dopants Al_{Zn} and Zn_{Al} (cat-S), where almost perfect agreement between the two functionals is achieved for $\Delta H_{\rm ref}^{GW,\rm FERE}$. The largest remaining deviations occur for the vacancy defects, indicating that this type of defect experiences the largest total-energy contribution [first term in Eq. (2)] due to the inclusion of Fock exchange in the hybrid functional. The complete list of defect formation energies is given in Table IV. As expected, we observe that the magnitude of the GW corrections $q\delta E_{\rm VBM}$ is much larger in GGA($+U$) (RMS = 2.32 eV) than in HSE (RMS = 0.93 eV), whereas the FERE corrections $\sum n_a \delta\mu_a$ are only somewhat smaller in HSE (RMS = 0.42 and 0.25, respectively). While the adjustment of the α parameter in the HSE functional to match the correct band gap would further reduce the magnitude of the GW band edge shifts and thereby improve the predictions over the standard $\alpha = 0.25$ value, the present approach removes the above noted ambiguities associated with such a materials-specific adjustment.

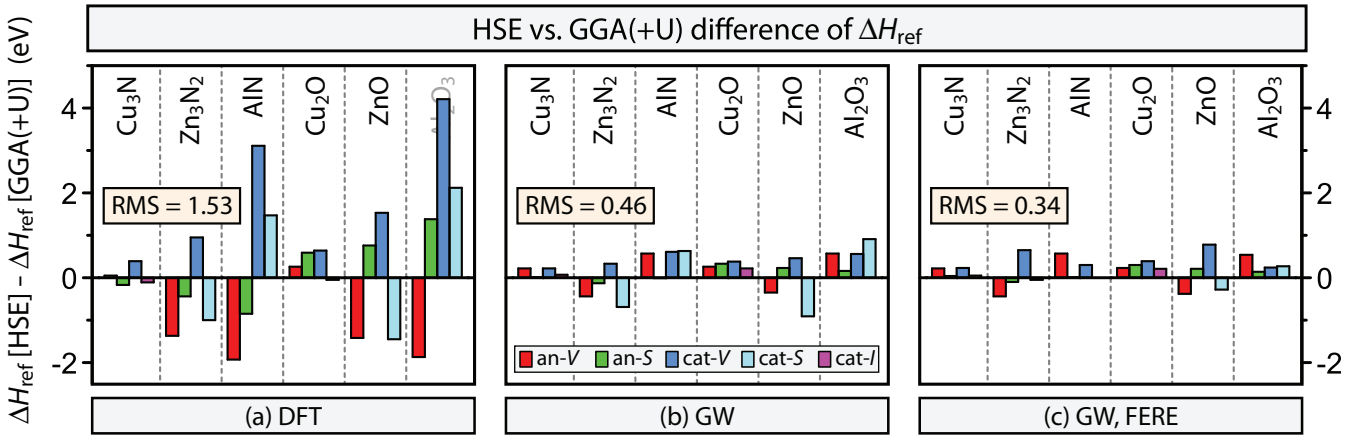

FIG. 2. (Color online) Difference of defect formation energies between the HSE and the GGA($+U$) functionals, based on (a) $\Delta H_{\rm ref}$, directly calculated in the respective DFT functional, on (b) $\Delta H_{\rm ref}^{GW}$, taking into account GW band edge shifts, and on (c) $\Delta H_{\rm ref}^{GW,\rm FERE}$, taking additionally into account the FERE shifts for the elemental reference energies [cf. Eq. (6)]. The different types of defects are anion vacancies (an-V: $V_{\rm N}$, $V_{\rm O}$), anion substitutional dopants (an-S: $O_{\rm N}$, $N_{\rm O}$), cation vacancies (cat-V: $V_{\rm Cu}$, $V_{\rm Zn}$, $V_{\rm Al}$), cation substitutional dopants (cat-S: $Al_{\rm Zn}$, $Zn_{\rm Al}$), and cation interstitials (cat-I: $Cu_{\rm i}$).

TABLE IV. The columns "GGA(+U)" and "HSE" give the defect formation energies ΔH_{ref}, as defined in Eq. (2), in the respective DFT Hamiltonian, and their values after correction for the quasiparticle energy shift ΔE_{VBM} determined from GW (see Table I) and after correction for the shifts $\delta\mu_a$ of the elemental reference energies determined in the FERE approach (see Table III). The column "ΔHSE-GGA(+U)" gives the respective differences between the two Hamiltonians. For the columns "GGA(+U)" and "HSE", the last row gives the root-mean-square (RMS) deviation of $\Delta H_{\text{ref}}^{\text{DFT}}$ and $\Delta H_{\text{ref}}^{GW}$ relative to the final result $\Delta H_{\text{ref}}^{GW,\text{FERE}}$. For the column "ΔHSE-GGA(+U)" it gives the RMS of the listed differences.

		GGA(+U)			HSE			ΔHSE-GGA(+U)		
	q	$\Delta H_{\text{ref}}^{\text{DFT}}$	$\Delta H_{\text{ref}}^{GW}$	$\Delta H_{\text{ref}}^{GW,\text{FERE}}$	$\Delta H_{\text{ref}}^{\text{DFT}}$	$\Delta H_{\text{ref}}^{GW}$	$\Delta H_{\text{ref}}^{GW,\text{FERE}}$	$\Delta H_{\text{ref}}^{\text{DFT}}$	$\Delta H_{\text{ref}}^{GW}$	$\Delta H_{\text{ref}}^{GW,\text{FERE}}$
Cu_3N										
V_N	+1	+0.75	+0.31	+0.28	+0.79	+0.53	+0.50	+0.05	+0.22	+0.22
O_N	+1	−1.14	−1.58	−1.95	−1.31	−1.57	−1.92	−0.17	+0.01	+0.04
V_{Cu}	−1	+1.10	+1.53	+1.52	+1.49	+1.75	+1.75	+0.39	+0.22	+0.23
Cu_i	+1	−0.04	−0.48	−0.46	−0.15	−0.41	−0.41	−0.11	+0.07	+0.05
Zn_3N_2										
V_N	+3	−0.53	−1.83	−1.86	−1.90	−2.27	−2.30	−1.37	−0.44	−0.44
O_N	+1	−3.46	−3.89	−4.26	−3.90	−4.02	−4.37	−0.44	−0.13	−0.10
V_{Zn}	−2	+2.65	+3.52	+3.44	+3.61	+3.85	+4.09	+0.95	+0.33	+0.65
Al_{Zn}	+1	−2.73	−3.16	−3.90	−3.73	−3.85	−3.95	−1.00	−0.69	−0.05
AlN										
V_N	+3	+2.28	−1.58	−1.61	+0.35	−1.01	−1.05	−1.93	+0.57	+0.57
O_N	+1	−4.10	−5.39	−5.76	−4.95	−5.40	−5.75	−0.85	−0.01	+0.01
V_{Al}	−3	+13.99	+17.85	+18.52	+17.10	+18.46	+18.81	+3.11	+0.61	+0.30
Zn_{Al}	−1	+4.60	+5.89	+6.63	+6.07	+6.52	+6.62	+1.47	+0.63	+0.00
Cu_2O										
V_O	0	+2.15	+2.15	+2.49	+2.50	+2.50	+2.82	+0.26	+0.26	+0.23
N_O	−1	+3.37	+4.04	+4.42	+3.94	+4.34	+4.69	+0.59	+0.33	+0.30
V_{Cu}	−1	+1.10	+1.77	+1.75	+1.74	+2.14	+2.14	+0.64	+0.38	+0.39
Cu_i	+1	+1.30	+0.64	+0.65	+1.28	+0.87	+0.88	−0.05	+0.22	+0.21
ZnO										
V_O	+2	+2.18	+0.00	+0.35	+0.76	−0.35	−0.03	−1.42	−0.35	−0.38
N_O	−1	+5.70	+6.78	+7.16	+6.46	+7.01	+7.36	+0.76	+0.23	+0.21
V_{Zn}	−2	+7.42	+9.59	+9.51	+8.95	+10.05	+10.29	+1.53	+0.46	+0.78
Al_{Zn}	+1	−4.37	−5.46	−6.20	−5.82	−6.37	−6.47	−1.45	−0.91	−0.28
Al_2O_3										
V_O	+2	+1.44	−2.40	−2.06	−0.43	−1.84	−1.52	−1.87	+0.57	+0.54
N_O	−1	+7.22	+9.14	+9.51	+8.59	+9.30	+9.65	+1.38	+0.16	+0.14
V_{Al}	−3	+14.95	+20.72	+21.38	+19.16	+21.27	+21.62	+4.21	+0.56	+0.24
Zn_{Al}	−1	+6.02	+7.94	+8.68	+8.14	+8.85	+8.95	+2.12	+0.91	+0.27
RMS-Δ		2.32	0.42		0.93	0.25		1.53	0.46	0.34

VI. CONCLUSIONS

In conclusion, by anchoring the charged-defect formation energy at the average electrostatic potential, we used the GW quasiparticle energy shifts of the band edge energies to address the bandgap problem in supercell defect calculations. Changes in the defect formation enthalpy of smaller but still significant magnitude result from the fitted elemental reference energies, which improve the thermochemical boundary conditions. Comparing results of density and hybrid functionals, we showed that both methods give very similar results for fully ionized defects when equivalent corrections are applied in either case. This finding suggests that accurate formation energies for the important case of fully ionized defects can be obtained even with standard GGA(+U) density functionals when using appropriate electronic and atomic reference energies. Existing results without such corrections can be easily reevaluated [see Eq. (6)]. For the hybrid functionals, this approach further provides an option to avoid using different exchange mixing parameters for different materials.

ACKNOWLEDGMENTS

Financial support: For NREL and CSM (H.P., V.S., and S.L.), the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Next Generation Photovoltaics II (SunShot initiative), under Contract No. DE-AC36-08GO28308 to NREL. For TCD (D.O.S and G.W.W.), the SFI through the PI programme (PI Grant Nos. 06/IN.1/I92 and 06/IN.1/I92/EC07). UCL (D.O.S.), Ramsay Memorial Trust and UCL Ramsay Fellowship. High performance computing resources: RedMesa by NREL's Computational Science Center, Kelvin by TCHPC, the Stokes cluster by ICHEC, and HECToR through membership of the HPC Materials Chemistry Consortium under

TABLE V. The elemental energies μ_a^{DFT} calculated in DFT and given for different PAW potentials of N and O along with the respective energy cutoff E_{cut}.

PAW potential	E_{cut} (eV)	$\mu_a^{\mathrm{GGA}(+U)}$ (eV)	μ_a^{HSE} (eV)
Cu ($U = 5$ eV in GGA)	320	−2.03	−4.08
Zn ($U = 6$ eV in GGA)	320	−0.56	−1.98
Al	320	−3.75	−4.22
N (soft)	320	−8.42	−11.04
N (standard)	500	−8.41	−11.55
O (soft)	320	−5.03	−7.95
O (standard)	500	−5.01	−8.63

EPSRC (Grant No. EP/F067496). Administrative support: For V.S., REMRSEC at CSM.

APPENDIX

Table IV summarizes the numerical results for the defect formation energies. The fully ionized state of anion-vacancies in Cu_2O and Cu_3N would theoretically be V_{O}^{2+} and V_{N}^{3+}, respectively, but these states are found to be unstable at $E_{\mathrm{F}} = E_{\mathrm{VBM}}$. Therefore, the energies for the stable V_{O}^{0} and V_{N}^{1+} states are given instead.

Computational details: The total energy and quasiparticle energy calculations were performed with the projector augmented wave (PAW) implementation of the VASP code, version 5.3.3, using the GGA (Al compounds), GGA + U (Cu, Zn compounds), and HSE functionals. The respective parameters were chosen as U(Cu-d) = 5 eV and U(Zn-d) = 6 eV for GGA + U, and as $\alpha = 0.25$, $\mu = 0.2$ Å^{-1} for the exchange-mixing and range-separation parameters in the HSE functional. The energy cutoff was varied between 320 and 500 eV as needed to reach convergence for the respective PAW potentials. Both the standard and soft versions for the PAW potentials of N and O were applied for total energy calculations. Due to the short bond distance in the N_2 and O_2 molecules (O_2 in spin-triplet state), the standard and soft PAW potentials do not give accurate energies for the N and O reference energies, i.e. $\mu_{\mathrm{N}}^{\mathrm{DFT}} = \frac{1}{2}E(\mathrm{N_2})$ and $\mu_{\mathrm{O}}{}^{\mathrm{DFT}} = \frac{1}{2}E(\mathrm{O_2})$. Therefore, we calculated the molecular binding energy with a hard PAW potential (900 eV cutoff), and then added half of the binding energy to atomic energy calculated with the standard or soft pseudopotential. The resulting energies relative to the nonspin-polarized spherical-symmetric free atom (according to the convention in VASP) are given in Table V. Differences in the order of 0.1 eV are observed with respect to a direct calculation of the molecules using the standard PAW potential (see Ref. 21). The defect calculations were performed in supercells of 72 atom or larger. A Γ-centered 2 × 2 × 2 k-mesh has been used for all supercell calculations. The HSE results were obtained with a high degree of convergence in the calculation of the exchange integral (VASP input settings NKRED = 1 and PRECFOCK = N). An explicit comparison of defect formation energies calculated with the standard and soft types of PAW potentials showed differences of less than 0.02 eV. Finite size effects for charged defects have been corrected according to Refs. 1 and 3, where the potential alignment has been determined by using all atoms except the defect site to calculate the offset of the average potential. For dielectric constant that enters into the finite size corrections, we used the total (electronic + ionic) low-frequency value $\varepsilon = \varepsilon_{\mathrm{elec}} + \varepsilon_{\mathrm{ion}}$. The electronic component $\varepsilon_{\mathrm{elec}}$ was determined from the dielectric matrix in the independent particle approximation in the respective GGA(+U) and HSE Hamiltonians, whereas the ionic part was taken from a density functional perturbation theory calculation in GGA(+U) for both cases. For Cu_3N and Zn_3N_2, we determined the equivalent finite-size correction terms also with the method of Ref. 2, finding that the potential alignment terms usually agree within about 20 meV. For the electrostatic image charge interaction, the method of Ref. 2 gave results closer to the full first-order term,[44] whereas the method of Refs. 1 and 3 typically give about 0.7 times the first-order term.

The GW calculations were performed on the relaxed GGA(+U) or HSE structures, where the unit cell vectors were rescaled to match the experimental lattice volume. The DFT wave functions were maintained, but the eigen energies were iterated to self-consistency. The response functions were calculated in the random phase approximation, and the local-field effects in the adiabatic LDA of time-dependent DFT were included. The soft PAW potentials were employed for N and O, with a cutoff of 320 eV for the wave functions and of 250 eV for the response functions. A total number of 64 bands per atom in the unit cell were used. For Cu and Zn compounds, and on-site potential V_d(Cu) = −2.4 eV and V_d(Zn) = −1.5 eV was applied to correct for the too-high d-orbital energy in GW (see Refs. 25 and 43). Slight differences compared to these previous works are due to improved convergence parameters for the exchange and the response functions.

[1] S. Lany and A. Zunger, Phys. Rev. B **78**, 235104 (2008).
[2] C. Freysoldt, J. Neugebauer, and C. G. van de Walle, Phys. Rev. Lett. **102**, 016402 (2009).
[3] S. Lany and A. Zunger, Modelling Simul. Mater. Sci. Eng. **17**, 084002 (2009).
[4] R. Ramprasad, H. Zhu, P. Rinke, and M. Scheffler, Phys. Rev. Lett. **108**, 066404 (2012).
[5] D. West, Y. Y. Sun, and S. B. Zhang, Appl. Phys. Lett. **101**, 082105 (2012).
[6] S. B. Zhang, S. H. Wei, and A. Zunger, Phys. Rev. B **63**, 075205 (2001).
[7] A. Janotti and C. G. van de Walle, Phys. Rev. B **76**, 165202 (2007).
[8] C. W. M. Castleton, A. Höglund, and S. Mirbt, Phys. Rev. B **73**, 035215 (2006).
[9] J. Paier, R. Hirschl, M. Marsman, and G. Kresse, J. Chem. Phys. **122**, 234102 (2005).
[10] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero,

A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys.: Condens. Matter **21**, 395502 (2009).
[11]J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
[12]J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003); **124**, 219906 (2006).
[13]F. Oba, A. Togo, I. Tanaka, J. Paier, and G. Kresse, Phys. Rev. B **77**, 245202 (2008).
[14]S. J. Clark and J. Robertson, Appl. Phys. Lett. **94**, 022902 (2009).
[15]D. O. Scanlon, B. J. Morgan, G. W. Watson, and A. Walsh, Phys. Rev. Lett. **103**, 096405 (2009).
[16]P. Agoston, K. Albe, R. M. Nieminen, and M. J. Puska, Phys. Rev. Lett. **103**, 245501 (2009).
[17]A. Janotti, J. B. Varley, P. Rinke, N. Umezawa, G. Kresse, and C. G. Van de Walle, Phys. Rev. B **81**, 085212 (2010).
[18]J. Pohl and K. Albe, Phys. Rev. B **87**, 245203 (2013).
[19]L. Hedin, Phys. Rev. **139**, A796 (1965).
[20]S. Lany, Phys. Rev. B **78**, 245207 (2008).
[21]V. Stevanovic, S. Lany, X. Zhang, and A. Zunger, Phys. Rev. B **85**, 115104 (2012).
[22]S. B. Zhang, D. Tomanek, S. G. Louie, M. L. Cohen, and M. S. Hybertsen, Solid State Commun. **66**, 585 (1988).
[23]R. Shaltaf, G.-M. Rignanese, X. Gonze, F. Giustino, and A. Pasquarello, Phys. Rev. Lett. **100**, 186401 (2008).
[24]A. Alkauskas, P. Broqvist, and A. Pasquarello, Phys. Rev. Lett. **101**, 046405 (2008); Phys. Status Solidi **248**, 775 (2011).
[25]L. Y. Lim, S. Lany, Y. J. Chang, E. Rotenberg, A. Zunger, and M. F. Toney, Phys. Rev. B **86**, 235113 (2012).
[26]S. Lany and A. Zunger, Phys. Rev. B **80**, 085202 (2009).
[27]H. Peng and S. Lany, Phys. Rev. B **85**, 201202(R) (2012).
[28]J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).
[29]P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, *ibid*. **140**, A1133 (1965).
[30]P. Mori-Sanchez, A. J. Cohen, and W. Yang, Phys. Rev. Lett. **100**, 146401 (2008).
[31]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).
[32]L. Wang, T. Maxisch, and G. Ceder, Phys. Rev. B **73**, 195107 (2006).
[33]A. Jain, G. Hautier, S. P. Ong, C. J. Moore, C. C. Fischer, K. A. Persson, and G. Ceder, Phys. Rev. B **84**, 045115 (2011).
[34]J. C. Grossman, J. Chem. Phys. **117**, 1434 (2002).
[35]J. Harl and G. Kresse, Phys. Rev Lett. **103**, 056401 (2009).
[36]J. Yan, J. S. Hummelshøj, and J. K. Nørskov, Phys. Rev. B **87**, 075207 (2013).
[37]D. D. Wagman, W. H. Evans, V. B. Parker, R. H. Schumm, I. Halow, S. M. Bailey, K. L. Churney, and R. L. Nutall, J. Phys. Chem. Ref. Data **11**, Suppl. 2 (1982).
[38]P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[39]G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[40]M. Shishkin and G. Kresse, Phys. Rev. B **74**, 035101 (2006).
[41]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[42]S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, Phy. Rev. B **57**, 1505 (1998).
[43]S. Lany, Phys. Rev. B **87**, 085112 (2013).
[44]G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).
[45]P. Rinke, A. Janotti, M. Scheffler, and C. G. van de Walle, Phys. Rev. Lett. **102**, 026402 (2009).
[46]S. Lany and A. Zunger, Phys. Rev. B **81**, 113201 (2010).