Intrinsic point-defect equilibria in tetragonal ZrO_2: Density functional theory analysis with finite-temperature effects

Mostafa Youssef and Bilge Yildiz*

Laboratory for Electrochemical Interfaces, Department of Nuclear Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

(Received 3 April 2012; revised manuscript received 17 July 2012; published 9 October 2012)

We present a density functional theory (DFT) framework taking into account the finite temperature effects to quantitatively understand and predict charged defect equilibria in a metal oxide. Demonstration of this approach was performed on the technologically important tetragonal zirconium oxide, T-ZrO_2. We showed that phonon free energy and electronic entropy at finite temperatures add a nonnegligible contribution to the free energy of formation of the defects. Defect equilibria were conveniently cast in Kröger–Vink diagrams to facilitate realistic comparison with experiments. Consistent with experiments, our DFT-based results indicate the predominance of free electrons at low oxygen partial pressure ($P_{O_2} \leqslant 10^{-6}$ atm) and low temperature ($T \leqslant 1500$ K). In the same regime of P_{O_2} but at higher temperatures, we discovered that the neutral oxygen vacancies (F-centers) predominate. The nature of the predominant defect at high oxygen partial pressure has been a long-standing controversy in the experimental literature. Our results revealed this range to be dominated by the doubly charged oxygen vacancies at low temperatures ($T \leqslant 1500$ K) and free electrons at high temperatures. T-ZrO_2 was found to be hypostoichiometric over all ranges of T and P_{O_2}, mainly because of the doubly charged oxygen vacancies, which are responsible for inducing n-type conductivity via a self-doping effect. A range of 1.3 eV in the band gap of T-ZrO_2 starting from the middle of the gap toward the conduction band is accessible to the chemical potential of electrons (Fermi level) by varying T and P_{O_2} without extrinsic doping. The approach presented here can be used to determine the thermodynamic conditions that extremize certain desirable or undesirable defects to attain the optimal catalytic and electronic performance of oxides.

DOI: 10.1103/PhysRevB.86.144109 PACS number(s): 61.72.J−, 71.55.−i, 65.40.gd, 31.15.es

I. INTRODUCTION

In 1926 Frenkel introduced the notion of point defects in crystalline solids to elucidate the phenomenon of diffusion.[1] Now we understand that not only diffusion but almost all properties of crystalline solids are affected by defects and, in particular, by point defects.[2] Metal oxides are no exception in terms of the dependence of their properties on the underlying point defects. This class of materials is currently widely used in various energy systems for a multitude of purposes,[3–6] and it is crucial to understand their defect equilibria and transport properties in order to advance these energy systems. While in operation, metal oxides are typically exposed to a wide range of temperatures and oxygen chemical potential, and this, in turn, renders our task of understanding and predicting their defect equilibria and transport properties even more difficult. In this paper we thoroughly analyze the defect equilibria in a technologically important metal oxide, namely tetragonal zirconium oxide (T-ZrO_2). Our approach is based on the density functional theory (DFT), taking into account finite temperature effects and spanning a wide range of oxygen chemical potentials. In a companion article we extend this analysis to include the transport kinetics of oxygen defects.[7] We believe that the framework we establish in these two articles serves as a model in analyzing point defects in metal oxides in realistic thermodynamic conditions.

Zirconia (ZrO_2) belongs to the list of the most important metal oxides due to its wide usage in different technological applications.[8] In energy systems it is used in fuel cells,[3] in gas sensors,[9] as a thermal barrier coating,[10] and as a protective natively grown layer against nuclear fuel cladding corrosion in light water nuclear reactors.[11–13] Its applications were recently extended to include usage as a gate dielectric for metal oxide semiconductor devices[5] and in biomedical applications such as hip implants[14] and dental restorations.[15] This diversity in zirconia applications is attributed to its superior properties; in particular it has high strength and fracture toughness,[16] an ionic conductivity that can be significantly enhanced by doping[17] and straining,[18] and high dielectric constant.[5] At ambient pressure zirconia crystallizes in three polymorphs. Up to 1440 K the monoclinic phase is stable; above this temperature the tetragonal phase is stable up to 2640 K, where it transforms to cubic phase; finally, melting occurs at 2870 K.[19] Among these three, the tetragonal phase has the highest fracture toughness[8,16] and the highest dielectric constant[20] and is the corrosion resistive phase.[11] Although tetragonal zirconia is thermodynamically stable at very high temperatures, it can be stabilized at much lower temperatures by stresses, doping, or reducing the crystallite size.[8] Improving the corrosion resistance,[11,12,21] enhancing the performance as a gate dielectric,[5] and preventing the well-known low-temperature degradation of T-ZrO_2[16] depend in the first place on our understanding of the defect physics in this phase.

Conductivity and thermogravimetry measurements were performed to understand the defect equilibria in T-ZrO_2.[22–26] However, these measurements could not fully resolve the nature of the native dominant point defects mainly due to the interfering role of impurities and the difficulty in performing the measurements at the very high temperatures in which this phase of zirconia is stable. In spite of this, all these experiments confirmed that at low oxygen partial pressure, T-ZrO_2 is a mixed ionic (predominantly with doubly charged oxygen

vacancies) and electronic (n-type) conductor, and at high oxygen partial pressure it is mainly an ionic conductor. Also, two experiments confirmed that the electronic transference number increases by elevating the temperature.[24,25] However, the type and the charge state of the predominant point defect at high oxygen partial pressure remained elusive. Furthermore, all the models proposed to understand the experimental measurements were based on fully ionized defects ignoring the ability of zirconia point defects to trap electronic charge carriers and hence changing their charge states. On the other hand, theoretical *ab initio* calculations focused only on oxygen vacancies,[27–29] and indeed only the study by Eichler[27] considered the charge states of this particular defect. Moreover none of the theoretical studies considered finite temperature effects or the role played by oxygen chemical potential on defect equilibria. Hence, a comparison between theory and experiments was never possible to perform.

We establish in this paper a framework to predict and understand defect equilibria in metal oxides, taking T-ZrO_2 as a model system. Our approach explicitly considers the effect of the relevant thermodynamic conditions in order to bridge the gap between theory and experiments. In particular we used DFT to calculate the formation energies and identify the defect structure for oxygen and zirconium defects in all possible charge states. Extension to free energies from DFT-calculated formation energies was accomplished by sampling the relevant excitations, mainly the phonon free energy and the electronic entropy. Finally, the overall picture of the defect equilibria was conveniently cast in a Kröger–Vink diagram. This was followed by a detailed comparison with the earlier reported experimental measurements. Our results are in agreement with the experimental finding that at low oxygen partial pressure ($P_{O_2} \leqslant 10^{-6}$ atm), free electrons predominate. In the same range of P_{O_2}, but at higher temperatures, we found that neutral oxygen vacancies (F-centers) are predominant. Moreover, our calculations revealed that at high P_{O_2}, the doubly charged oxygen vacancies predominate at low temperatures ($T \leqslant 1500$ K) and free electrons predominate at high temperatures. We extended our analysis to track the dependence of two important variables, namely the chemical potential of electrons (Fermi level) and off-stoichiometry on the thermodynamic conditions. We found that T-ZrO_2 is hypostoichiometric in all ranges of T and P_{O_2}. Additionally, a range of 1.3 eV in the band gap of T-ZrO_2 from the middle of the gap toward the conduction band is accessible to the chemical potential of electrons by a self-doping effect from the doubly charged oxygen vacancies, which cause n-type conductivity. In a companion paper,[7] we use the defect equilibria established here as the backbone to predict oxygen transport kinetics in T-ZrO_2 in the wide range of thermodynamic conditions explored here.

II. THEORETICAL AND COMPUTATIONAL APPROACH

This section is presented in four subsections. First, we present the relevant thermodynamics of defect formation in a metal oxide. Second, we describe the algorithm of constructing the Kröger–Vink diagram. Third, we explain how we included finite temperature effects in our calculations. Finally, we describe the needed DFT calculations.

A. Thermodynamics of defects formation

To understand the energetics of the formation of a point defect in a metal oxide, exemplified here by tetragonal ZrO_2, we appeal to the picture of a constrained grand canonical ensemble. In this picture a crystal of ZrO_2 of a fixed volume is constrained to exchange oxygen, but not zirconium, with an oxygen reservoir of a fixed temperature and oxygen chemical potential. In fact the volume does not have to be fixed, but we assume that the change of volume due to the creation of point defects is negligible when these defects have dilute concentrations. In this constrained ensemble, physically, the creation of all types of point defects occurs by exchanging only oxygen with the reservoir. Thus an oxygen vacancy forms by removing an oxygen atom from the crystal and inserting it in the reservoir. Conversely, incorporating an oxygen atom from the reservoir into an interstitial site in the crystal leads to the formation of an interstitial oxygen. To create a zirconium vacancy, two oxygen atoms get incorporated in the crystal in regular lattice sites representing a ZrO_2 unit formula missing the zirconium cation. On the other hand creating a zirconium interstitial requires moving two oxygen atoms from the crystal into the reservoir and displacing a zirconium cation from its original lattice site to an interstitial site. In all these reactions charge neutrality has to be maintained. We elaborate on the latter condition in the next subsection. While this formalism does not rely on the traditional concepts of Frenkel and Schottky disorder, accounting for them is still amenable within this constrained ensemble, as can be readily understood on the basis of the defect creation processes described above (simply by composing Frenkel or Schottky from their constituents) and as will be shown quantitatively later in this subsection. The picture described above is necessary to understand the energetics of point defect creation, however, the actual modeling of these defects in a DFT supercell is described in Sec. II D. We can now define the Gibbs free energy of formation, $G^{\mathrm{f}}_{D,q}$, of defect D that has a charge q as follows,

$$G^{\mathrm{f}}_{D,q} = \Delta E^{\mathrm{solid}}_{\mathrm{potential}} + \Delta E^{\mathrm{solid}}_{\mathrm{vib}} - T\left(\Delta S^{\mathrm{solid}}_{\mathrm{vib}} + \Delta S^{\mathrm{solid}}_{\mathrm{elec}}\right) + P\Delta V \pm \mu_D, \quad (1)$$

where $\Delta E^{\mathrm{solid}}_{\mathrm{potential}}$, $\Delta E^{\mathrm{solid}}_{\mathrm{vib}}$, $\Delta S^{\mathrm{solid}}_{\mathrm{elec}}$, $\Delta S^{\mathrm{solid}}_{\mathrm{vib}}$, and ΔV are the differences in potential energy, vibrational energy, electronic entropy, vibrational entropy, and volume between the defected and perfect crystals. P is the pressure, T is the temperature, and μ_D is the chemical potential of species D. As discussed above the change in volume is negligible; hence, the $P\Delta V$ term can be set to zero. The term $\Delta E^{\mathrm{solid}}_{\mathrm{vib}} - T\Delta S^{\mathrm{solid}}_{\mathrm{vib}} = F^{\mathrm{solid}}_{\mathrm{vib}}$ represents the phonon contribution to the free energy of formation of the defect. This contribution is significantly enhanced with temperature. The term $-T\Delta S^{\mathrm{solid}}_{\mathrm{elec}}$ is also enhanced with the temperature and adds a nonnegligible contribution to the free energy of formation of the defect. Subsection II.C addresses these finite temperature effects. It is worth noting that the convention is to exclude the configurational entropy of the defects from the definition of the Gibbs free energy. However, the contribution of this entropy is still accounted for implicitly when computing the concentrations of the defects, as described in the next subsection. Indeed defects are stabilized in crystals at finite temperature mainly due to the configurational entropy.[30] Finally, we define the term

$\Delta E^{\text{solid}}_{\text{potential}} \pm \mu_D$ as the formation energy of the defect and denote it by $E^{\text{f}}_{D,q}$. Next we describe the details of the latter term.

The formation energies of oxygen and zirconium defects can be computed as

$$E^{\text{f}}_{\text{O},q} = E_{\text{defected}} - E_{\text{perfect}} \pm \tfrac{1}{2}\mu_{\text{O}_2} + q(E_{\text{VBM}} + \mu_{\text{F}}) + E_{\text{MP}}, \tag{2}$$

$$E^{\text{f}}_{\text{Zr},q} = E_{\text{defected}} - \frac{N \mp 3}{N} E_{\text{perfect}} \mp \mu_{\text{O}_2} + q(E_{\text{VBM}} + \mu_{\text{F}}) + E_{\text{MP}}, \tag{3}$$

where the top signs are for vacancies and the bottom signs are for interstitials. E_{defected} is the DFT energy of the supercell that contains the defect. E_{perfect} is the DFT energy of the perfect supercell. μ_{O_2} is the chemical potential of oxygen. E_{VBM} is the energy of the valence band maximum in the perfect supercell. μ_{F} is the chemical potential of electrons (or the Fermi level) relative to the valence band maximum; hence, it can take values from 0 to the width of the band gap in the perfect crystal. E_{MP} is the Makov–Payne correction for the multipole interactions between the periodic images of the charged defects.[31] Further details about the computation of formation energies and the accuracy of the Makov–Payne correction are given in the Supplemental Material.[32]

The chemical potential of oxygen, μ_{O_2}, is defined as[33]

$$\mu_{\text{O}_2}(T, p_{\text{O}_2}) = E^{\text{DFT}}_{\text{O}_2} + \mu^0_{\text{O}_2}(T, P^0) + k_{\text{B}}T \ln\left(\frac{p_{\text{O}_2}}{P^0}\right), \tag{4}$$

where $E^{\text{DFT}}_{\text{O}_2}$ is the total energy of an isolated oxygen molecule computed by DFT, p_{O_2} is the partial pressure of the oxygen gas, and k_{B} is the Boltzmann constant. $\mu^0_{\text{O}_2}(T, P^0)$ is the difference in chemical potential of O_2 between $T = 0$ K and the temperature of interest at the reference pressure P^0, which is typically taken as 1 atm. The term $\mu^0_{\text{O}_2}(T, P^0)$ was obtained from thermo-chemical tables.[34] It is important to note that Eq. (3) implicitly indicates that the chemical potential of zirconium, μ_{Zr}, is defined as

$$\mu_{\text{zr}} = E^{\text{DFT}}_{\text{ZrO}_2} - \mu_{\text{O}_2}, \tag{5}$$

where $E^{\text{DFT}}_{\text{ZrO}_2}$ is the DFT energy of the unit formula of ZrO_2 in the perfect crystal of zirconia. This definition is consistent with the constrained grand canonical ensemble picture that we described in the beginning of this section and has been adopted in studying charged defects in other oxides.[35,36] As mentioned in the beginning of this subsection, it is still possible to account for the traditional concepts of Frenkel and Schottky disorder. The formation energy of composite defects such as the Frenkel pair or Schottky defect becomes the sum of the individual formation energies of the components of the composite defects, as can be readily seen from Eqs. (2) and (3).

B. Construction of the Kröger–Vink diagram

The basic idea that governs the Kröger–Vink diagram[30] is that at a given temperature and oxygen partial pressure, the concentration of ionic and electronic defects should sum up to achieve charge neutrality in the bulk of the metal oxide. This statement can be expressed in mathematical terms as

$$\sum_{D,q} q[D^q] + p_{\text{v}} - n_{\text{c}} = 0, \tag{6}$$

where $[D^q]$ is the concentration of a point defect D of charge q, p_{v} is the concentration of holes in the valence band, n_{c} is the concentration of electrons in the conduction band, and the summation is taken over all charged defects. The concentration of a point defect D with charge q was derived by Kasamatsu *et al.*[37] by minimizing the Gibbs free energy of the defected crystal and hence accounting for the configurational entropy. The expression is

$$[D^q] = n_D \frac{\exp\left(-\frac{G^{\text{f}}_{D,q}}{k_{\text{B}}T}\right)}{1 + \sum_{q'} \exp\left(-\frac{G^{\text{f}}_{D,q'}}{k_{\text{B}}T}\right)}, \tag{7}$$

where n_D is the number of possible sites for the defect D in the lattice per chemical formula of ZrO_2; thus, the concentration is given in units of number of defects per chemical formula. The functional form of the concentration of a point defect is similar to the Fermi–Dirac distribution, although no quantum mechanical considerations were accounted for in the derivation. The assumption that $G^{\text{f}}_{D,q'} \gg k_{\text{B}}T$ for all charges q' leads to neglecting the summation in the denominator and obtaining the familiar Boltzmann-like expression for a point defect concentration. However, we are not adopting this assumption as it breaks down at high temperature, at extreme oxygen partial pressure conditions, or both. Special care needs to be taken in applying the summation in Eq. (7). This is a summation over the charge states of the defect that compete for the same site in the lattice. (See the Supplemental Material[32] for an illustration.)

The concentrations of electrons in the conduction band and holes in the valence band are given by applying Fermi–Dirac statistics to the electronic density of states (DOS)[38] as follows,

$$n_{\text{c}} = \int_{E_{\text{CBM}}}^{\infty} g_{\text{c}}(E) \frac{dE}{1 + \exp\left(\frac{E - \mu_{\text{F}}}{k_{\text{B}}T}\right)}, \tag{8}$$

$$p_{\text{v}} = \int_{-\infty}^{E_{\text{VBM}}} g_{\text{v}}(E) \frac{dE}{1 + \exp\left(\frac{\mu_{\text{F}} - E}{k_{\text{B}}T}\right)}, \tag{9}$$

where $g_{\text{c}}(E)$ and $g_{\text{v}}(E)$ are the density of electronic states in the conduction band and valence band per volume of the chemical formula ZrO_2, respectively. E_{CBM} is the energy of the conduction band minimum.

With the above expressions, the construction of the diagram at a fixed temperature proceeds by examining a wide range of the oxygen partial pressure. At each value of the latter, there is only one value for the chemical potential of electrons, μ_{F}, that achieves the charge neutrality condition. This value of the electron chemical potential has to satisfy two constraints; otherwise, we designate the oxygen partial pressure under examination as inaccessible. The first constraint is that the electron chemical potential has to have a value between the valence band maximum and the conduction band minimum. The second is that the formation energies, $E^{\text{f}}_{D,q}$, of all the defects have to be positive because a negative value of the formation energy indicates that the crystal is not stable at these

values of the electron chemical potential, and a phase transition or phase separation would take place at such values to avoid instabilities in the crystal. A final constraint that we applied to the overall construction of the Kröger–Vink diagram is limiting the search for accessible thermodynamic conditions (T, P_{O_2}) to an off-stoichiometry, x, in ZrO_{2+x}, in the range $|x| < 0.065$. These extrema of the values of x correspond to the extrema of the actual off-stoichiometry modeled in our DFT supercells. We regard this range of x as the range in which the assumption of noninteracting defects holds. Typical off-stoichiometries of undoped T-ZrO_2 fall within the limit of noninteracting defects that we specified. On the other hand, this does not hold for other metal oxides that exhibit higher off-stoichiometries, such as CeO_2 and UO_2.

It should be emphasized at this stage that the temperature, the oxygen partial pressure, and the electron chemical potential are not independent variables. Indeed, specifying two of them is, in principle, enough to determine the third.

C. Finite temperature effects

A major challenge that confronts bringing the DFT calculations (at 0 K) closer to the realm of experimental conditions is capturing the finite temperature effects. Ignoring these effects can lead to results that are far from being quantitative.[35,39] The major difficulty in capturing these effects is that they are typically very expensive computationally. In our modeling of defect equilibria in tetragonal zirconia we found that the free energy of phonons and the electronic entropy have important quantitative contributions. This is the case especially because the undoped T-ZrO_2 is stable at very high temperatures. In this subsection we examine these finite temperature effects.

We illustrate these finite temperature effects on the neutral oxygen vacancy in T-ZrO_2 as an example here. (The same approach and analysis are applicable to all the other defects.) The creation of this defect is associated with a change in the vibrational free energy of the system, ΔF_{vib}, which can be written as

$$\Delta F_{\text{vib}} = F_{\text{vib}}^{\text{defected}} + F_{\text{vib}}^{\text{O}} - F_{\text{vib}}^{\text{perfect}}. \tag{10}$$

The vibrational free energy of the oxygen atom in the gas phase, $F_{\text{vib}}^{\text{O}}$, is already included in the chemical potential term for the oxygen atom in the gas phase, μ_{O}. What is typically neglected in defect calculations are the vibrational free energies of the perfect crystal, $F_{\text{vib}}^{\text{perfect}}$, and the defected crystal, $F_{\text{vib}}^{\text{defected}}$. We accounted for these two terms within the harmonic approximation[38] as follows:

$$F_{\text{vib}}^{\text{solid}} = N k_{\text{B}} T \int_0^{\infty} g(\omega) \ln\left[2 \sinh\left(\frac{h\omega}{4\pi k_{\text{B}} T}\right)\right] d\omega. \tag{11}$$

N is the number of degrees of freedom, ω is the phonon frequency, $g(\omega)$ is the phonon DOS, and h is Planck's constant.

Similarly a change in the electronic entropy occurs in the system upon creating a neutral oxygen vacancy as follows:

$$\Delta S_{\text{elec}} = S_{\text{elec}}^{\text{defected}} + S_{\text{elec}}^{\text{O}} - S_{\text{elec}}^{\text{perfect}} \tag{12}$$

The electronic entropy for the oxygen atom in the gas phase, $S_{\text{elec}}^{\text{O}}$, is already accounted for in the chemical potential term for the oxygen atom, μ_{O}. However, it is common to neglect the electronic entropies of the perfect crystal, $S_{\text{elec}}^{\text{perfect}}$, and the defected crystal, $S_{\text{elec}}^{\text{defected}}$. In fact the electronic entropy for a wide–band gap perfect crystal is negligible. However, if a point defect was introduced in the crystal and this defect subsequently introduced electronic states in the band gap, then the electronic entropy contribution of these states can be significant.[40] The electronic entropy for a solid crystal was calculated as[38,41]

$$S_{\text{elec}}^{\text{solid}} = -k_{\text{B}} \int_{-\infty}^{\infty} g(E)[f \ln f + (1-f)\ln(1-f)] dE, \tag{13}$$

where $g(E)$ is the electronic DOS and f is the Fermi–Dirac distribution, which is a function of temperature, energy, and the chemical potential of electrons. It is worth noting that all the temperature dependence of $S_{\text{elec}}^{\text{solid}}$ is embedded in the Fermi–Dirac distribution since we neglected the temperature dependence of the electronic DOS. An accurate determination of the temperature dependence of $g(E)$ is still beyond the current computational capabilities. (See the Supplemental Material[32] for a discussion on the challenges of modeling the temperature dependence of $g(E)$.) However, in the Results and Discussion section we present Kröger–Vink diagrams calculated at different band gaps, thus indicating how the reduction of the band gap, which is the main impact of temperature on $g(E)$,[42] can alter the defect equilibria.

In Fig. 1(a) we plotted the correction to the free energy of formation of the neutral oxygen vacancy due to the phonon vibrational contribution in the solid ($F_{\text{vib}}^{\text{defected}} - F_{\text{vib}}^{\text{perfect}}$) and the electronic entropy contribution in the solid ($-T(S_{\text{elec}}^{\text{defected}} - S_{\text{elec}}^{\text{perfect}})$). Both the temperature and oxygen partial pressure dependence of these corrections are also shown in Fig. 1. It is clear that phonons add a significant quantitative correction to the free energy of formation, especially at the very high temperatures. This correction is further amplified by the exponential dependence of the defect concentration on the free energy of formation. Within the assumptions adopted here, the phonon correction is independent of the variation of P_{O_2} on an isotherm. Quantitatively the correction due to the change of the electronic entropy of the solid is not as significant as that of phonons. However on an isotherm, this correction exhibits an interesting dependence on P_{O_2} that originates from the $-T S_{\text{elec}}^{\text{defected}}$ term. As will be shown in the Results and Discussion section, at $P_{O_2} = 1$ atm the chemical potential of electrons μ_{F} is located almost at the middle of the band gap. Then by lowering P_{O_2}, it starts moving up toward the conduction band edge. Once μ_{F} reaches a value of about 2.6 eV (i.e., coincident with the defect state shown in Fig. 1(b)) it resonates with a defect-induced state in the band gap, and as a result the electronic entropy is enhanced, and this leads to the minima observed in Fig. 1(a). The role of electronic entropy in defect equilibria is not well studied, particularly in nonmetal systems. The insights that we demonstrated here conveys an important message that whenever a defect state arises in the band gap within the reach of μ_{F}, the electronic entropy contribution should be assessed carefully.

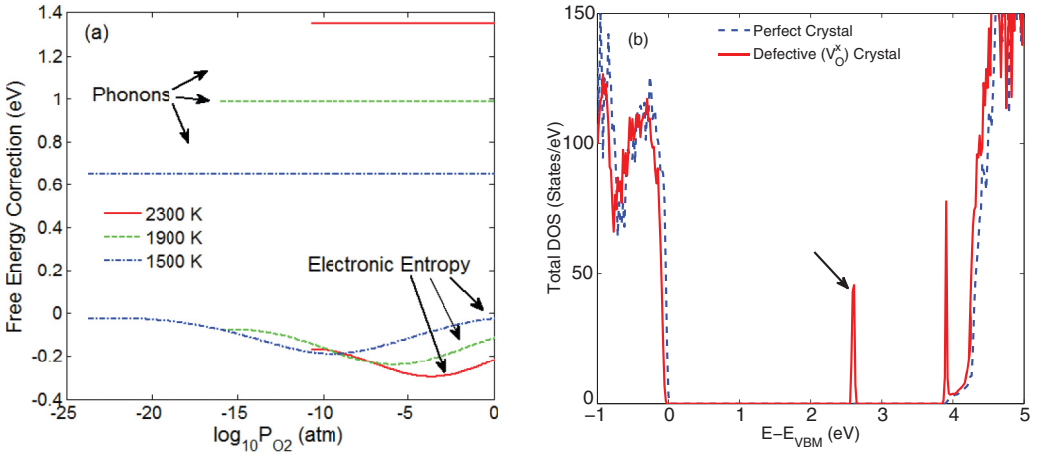

FIG. 1. (Color online) (a) Corrections to the free energy of formation of the neutral oxygen vacancy due to the change in phonon vibrations and the electronic entropy in the solid. (b) The total electronic DOS for tetragonal ZrO_2 in a perfect supercell and in a supercell with a neutral oxygen vacancy. The arrow in (b) indicates the electronic state in the gap due to the neutral oxygen vacancy that enhances the electronic entropy of the defect. The position and magnitude of the resonance [in (a)] between μ_F and the defect state [in (b)] depends on temperature and oxygen partial pressure.

D. Density functional theory calculations

Density functional theory calculations were performed using the projector-augmented plane-wave method[43] as implemented in the Vienna *Ab initio* Simulation Package (VASP).[44–47] Exchange correlation was treated in the generalized gradient approximation (GGA) as parameterized by Perdew, Burke, and Ernzerhof (PBE).[48,49] In this work we chose to apply standard PBE instead of a hybrid functional that incorporates a fraction of the exact exchange interaction for the following reasons. First, the typical practice of using hybrids requires a precise knowledge of the band gap in order to choose a fraction of the exact exchange that reproduces the previously known band gap. In the case of T-ZrO_2, there are only two reported values for the experimental band gap: 4.2 and 5.7 eV.[50,51] Those two values are too widely separated from each other to be useful while choosing the exact exchange fraction. Second, zirconium oxide is not among the families of strongly correlated metal oxides;[52] hence, standard PBE is able to describe this system reasonably. Third, and most important, the picture that hybrid functionals can act as *the* norm, to which the errors in standard PBE predictions can be referenced, has been challenged recently by Ramprasad *et al.*[53] Those authors demonstrated that while standard PBE underestimates quantities, such as the valence band width and the formation energies with respect to accurate experimental determination, hybrid functionals overestimate these quantities. Fourth, a major goal in the current work is to emphasize the contribution of finite temperature effects (in particular phonons and electronic entropy) to the defect equilibria in metal oxides. At the temperatures of interest in this work, these effects can be more dominant compared with any error due to standard PBE band gap issues. Additionally, evaluating these finite temperature effects using hybrid functionals is too expensive computationally and is not anticipated to be more accurate than standard PBE because they tend to overestimate the vibrational frequencies.[54] Finally, in order to address the underestimation of the band gap due to PBE, we performed a sensitivity analysis on the predicted Kröger–Vink diagram, as discussed in the Results and Discussion section.

For zirconium the $4s^2\ 4p^6\ 4d^2\ 5s^2$ electrons were treated as valence electrons. For oxygen the $2s^2\ 2p^4$ electrons were treated as valence electrons. These electrons were represented by a set of plane waves expanded up to a kinetic energy cutoff of 450 eV. The error in the total energy with respect to calculations performed using a kinetic energy cutoff of 600 eV was found to be less than 5 meV per chemical formula. Bulk point defect calculations were performed in a supercell that is $2 \times 2 \times 2$ of the conventional unit cell (see the Supplemental Material[32] for perfect crystal lattice parameters). The later has four zirconium cations and eight oxygen anions. In this paper all crystallographic directions are based on the conventional unit cell. A $2 \times 2 \times 2$ Monkhorst–Pack k-point mesh was used to perform reciprocal space integrals. The error in the total energy with respect to calculations performed with a $6 \times 6 \times 6$ k-point mesh was found to be less than 1 meV per chemical formula. Gaussian smearing with a smearing width of 0.05 eV was used to accelerate the convergence of the electronic structure. All ionic relaxations were considered converged when the forces on all ions were less than 0.01 eV/Å. For computational efficiency, all calculations were done initially by seeking a net zero magnetic moment solution, and after convergence spin-polarized calculations were performed on the preconverged structure.

Single defects were introduced in the super cell one a time. No defect-defect association was considered as the equilibrium concentrations of all defects in undoped tetragonal zirconia are expected to be dilute. We considered oxygen vacancies with charge states from 0 to $2+$, oxygen interstitials with charges from $2-$ to 0, zirconium vacancies with charges from $4-$ to 0, and zirconium interstitials with charges from 0 to $4+$. Antisites were not considered here as they are expected to have high formation energies, in particular for zirconia[55] and in general for all binary ionic materials due to electrostatic considerations.[2]

The zirconium sublattice in T-ZrO_2 is distorted, face centered cubic and hence is close-packed. Therefore, we considered only interstitial zirconium at the octahedral site in the center of the conventional unit cell. On the other hand, the oxygen sublattice is more open as it is a distortion for the simple cubic structure. Thus, we considered eight different interstitial structures. These are ⟨100⟩, ⟨001⟩, ⟨110⟩, ⟨101⟩, and ⟨111⟩ split dumbbells; ⟨100⟩ and ⟨001⟩ crowdions; and the octahedral site in the center of the conventional unit cell. This nomenclature of interstitial structures is borrowed from the literature of defects in metals. This comprehensive examination of interstitial configurations, however, has not been common in studying defects in ionic materials. In this paper we chose to transfer this nomenclature to ionic materials to emphasize the importance of exhausting many possible interstitial configurations when searching for the one with lowest energy. Figure 2 exemplifies representative interstitial structures for oxygen in T-ZrO_2.

For all the charge states that we considered, the interstitial oxygen in the ⟨100⟩ crowdion structure significantly distorted the tetragonal structure of zirconia in a manner that makes it difficult to define its formation energy with respect to the perfect crystal T-ZrO_2. The ⟨100⟩ crowdion is not considered further in this paper.

DFT calculations were also performed to account for phonons within the harmonic approximation. According to the symmetry of the supercell, certain atoms in certain directions were displaced twice, backward and forward, a distance of 0.004 Å, and the electronic ground state was calculated. This allowed, utilizing a central finite difference, the construction of the Hessian matrix and the determination of the vibrational frequencies of the system. The code PHONOPY[57] was then used to calculate the phonon DOS using 15 × 15 × 15 k-point sampling centered at the gamma point.

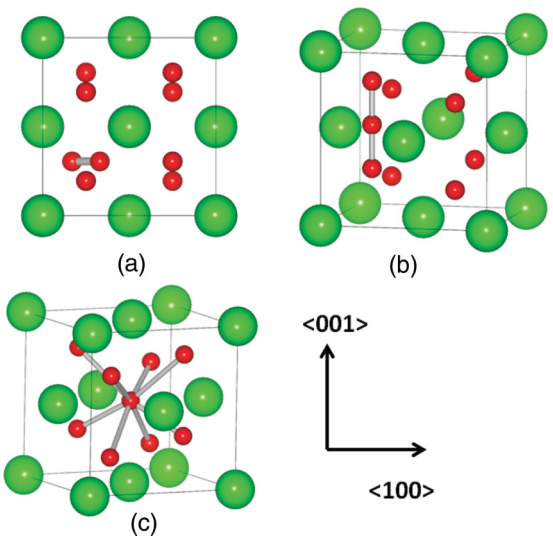

FIG. 2. (Color online) Examples of the oxygen interstitial structures considered: (a) ⟨110⟩ dumbbell, (b) ⟨001⟩ crowdion, (c) the octahedral site. Green (large) and red (small) balls represent zirconium and oxygen, respectively. The sticks are a guide for the eye and have no particular physical meaning. This figure was generated using the visualization software VESTA.[56]

To calculate the concentration of free electrons and holes (as described in subsection II.B) and to quantify the contribution of electronic entropy to the free energy of formation of the point defects, the electronic DOS were calculated for the perfect and defected T-ZrO_2 supercells. In these calculations a finer k-point mesh of 7 × 7 × 7 centered at the Γ point was used in combination with the tetrahedron method with Blöchl corrections.[58]

III. RESULTS AND DISCUSSION

This section is presented in four subsections. In the first subsection we discuss the dependence of the formation energies of the charged point defects on the electron chemical potential. In the second subsection, we present our calculated Kröger–Vink diagrams and compare them with conductivity measurements from the literature. In particular, the slopes of the calculated defect concentrations in the Kröger–Vink diagrams are found consistent with the slopes of the total conductance in the conductivity measurements from literature. Hence, our prediction of the predominance of free electrons at low P_{O_2} is consistent with prior experimental conclusions. A key contribution of our DFT-based results is the finding that the doubly charged oxygen vacancies are the dominant defects in the high P_{O_2} regime, whereas the prior experiments had fallen short of revealing the type and charge state of the predominant defect. In the third subsection we discuss the variations of T-ZrO_2 off-stoichiometry and electron chemical potential with the thermodynamic conditions. T-ZrO_2 was found to be hypostoichiometric in the entire T and P_{O_2} range. Furthermore, the chemical potential of electrons has access to a range of 1.3 eV of the band gap of T-ZrO_2 by a self-doping effect, mainly from the doubly charged oxygen vacancies. Finally, we discuss our key observations related to the atomic and electronic structure of the point defects.

A. Defect formation energies at 0 K

Figures 3(a)–3(c) is a plot of the formation energies at 0 K (excluding the zero point energy) for all the defects we considered as a function of the electron chemical potential. Figure 3(d) summarizes the thermodynamic transition levels between the dominant charge states of each defect as extracted from panels (a–c). Two important notes need to be clarified for this figure. First, the formation energies were allowed to take negative values. This is not be confused with the fact that the Gibbs free energy of a defective crystal is more negative than that of a perfect crystal at finite temperature,[2] since what we plotted in Fig. 3 is indeed the formation enthalpy of the defects neglecting the pressure-volume term. The enthalpy of formation for a point defect is a positive quantity,[2,30] and this was taken care of while constructing the Kröger–Vink diagram, as pointed out in the Theoretical and Computational Approach section. Second, the formation energies of oxygen interstitials were based on the energetically most favorable configurations (see Sec. III D).

As was shown previously for T-ZrO_2 by Eichler,[27] we found that the oxygen vacancies exhibit negative-U behavior, with a U value of −0.037 eV. This behavior indicates that the singly charged vacancies are not stable with respect to the disproportionation reaction into neutral and doubly charged

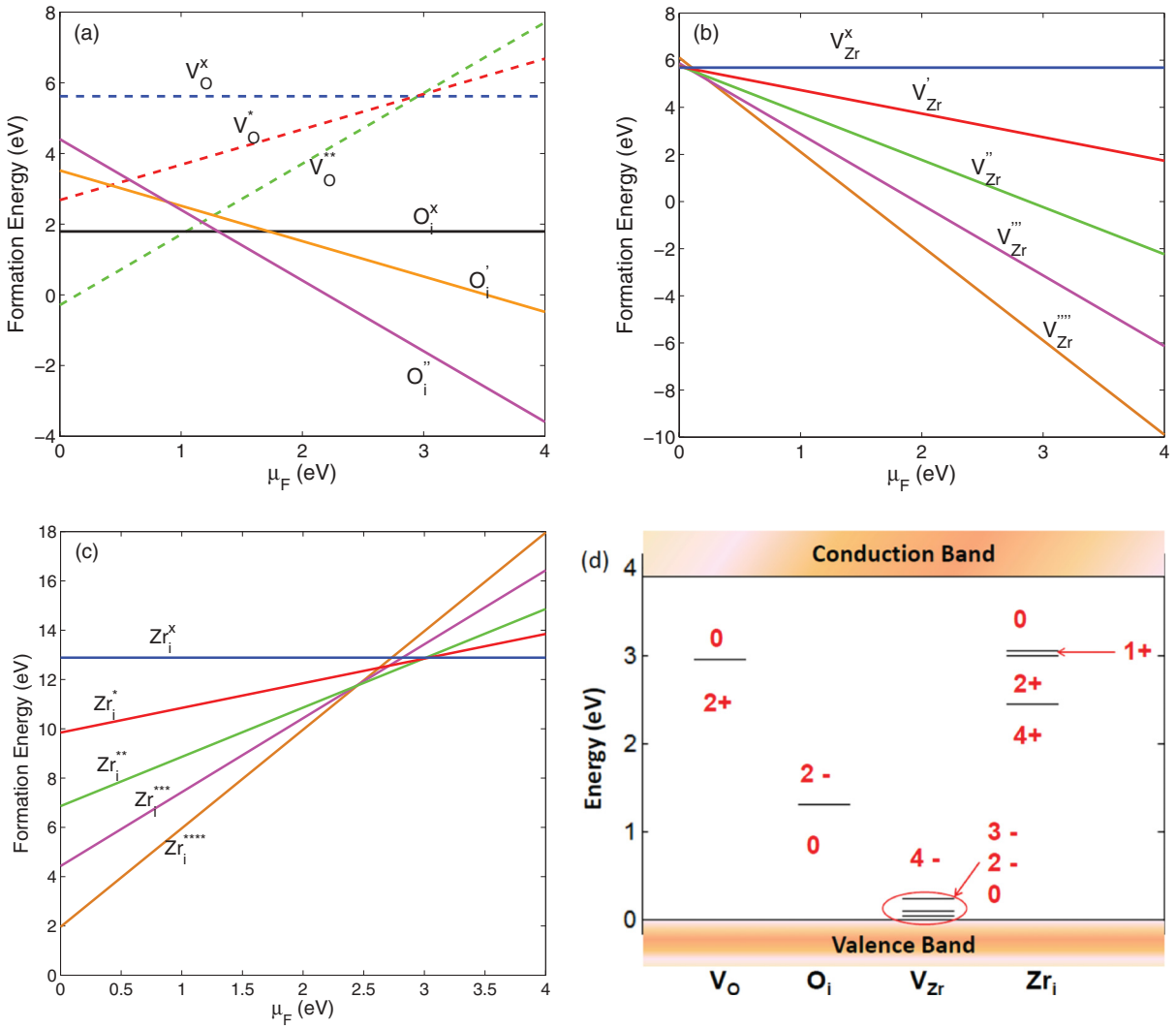

FIG. 3. (Color online) Defect formation energies as a function of the electron chemical potential. (a) Oxygen defects. (b) Zirconium vacancies. (d) Zirconium interstitials. (d) The thermodynamic transition levels for the native defects determined from panels (a–c).

vacancies. We also found that oxygen interstitials exhibit negative-U behavior, with a U value of -0.838 eV. Although the negative-U behavior indicates that the intermediate charge state is never stable thermodynamically, signatures of this banned charge state were observed in electron spin resonance spectra[59] in the context of studying hydrogen defects in wide–band gap oxides. This was explained by metastability due to sufficient isolation of the disallowed charged defect.[59] However, there is no transparent way to quantify the concentration of the defects that are in a metastable charge state, and in any case, it is expected that this concentration is less than what would be calculated on the basis of their formation energies. Thus, to simplify the analysis we will assume that the disallowed charge states, $V_O^{\bullet}$ and O_i', always have negligible concentration. If under any circumstances their concentrations increase to a nonnegligible level, once equilibrium is reached, the majority of these defects will undergo the disproportionation reaction; hence, their concentration drops to a very low level again.

The quadruply charged vacancy is the predominant zirconium vacancy throughout most of the band gap compared with the other charge states, as in Fig. 3(b). This can be understood on the basis that this is the only charge state for zirconium vacancies that does not lead to the formation of the O^- ion, as will be explained in subsection III.D. Although O^- is favorable to O^{2-} in the gas phase, the Madelung potential in an ionic crystal favors the O^{2-} ion;[60] thus, V_{Zr}'''' is predominant. On the other hand, zirconium interstitials have in general high formation energies compared with all other native defects, as shown in Fig. 3(c). Figure 3(d) depicts the thermodynamic transition levels of all the native defects. A thermodynamic transition level indicates the value of the electron chemical potential at which the dominant charge state of a defect changes from one value to another. These levels are amenable to experimental determination. Deep levels in the band gap can be measured by deep-level transient spectroscopy experiments, while shallow levels can be determined through temperature-dependent Hall measurements.[61] The diagram in Fig. 3(d) is more relevant to studying zirconia as a gate dielectric material, compared with the traditional Kröger–Vink diagram, which is more common in ceramics literature. Furthermore, the thermodynamic transition levels can correspond directly to a

separation between two regions in the Kröger–Vink diagram, as will be discussed in Sec. III B.

We turn now to compare our results of the formation energies with the values available in the literature. A detailed comparison is in the supplementary material.[32] In brief, our results for the oxygen vacancy formation energies are in good agreement with the DFT calculations of Eichler[27] and of Ganduglia-Pirovano *et al*.[28] We also have reasonable agreement with the embedded cluster Hartree–Fock and Becke–three-parameter–Lee–Yang–Parr (B3LYP)-DFT calculations of Safonov *et al*.[29] We are not aware of any published electronic structure calculations of defects other than oxygen vacancies in undoped tetragonal zirconium oxide. However, the study of Dwivedi *et al*.[62] using classical interionic potential considered all fully ionized defects of T-ZrO_2. The predictions of this potential are far from being reasonable, as is expected since a simple pair potential with a fixed charge for each ion is not able to reproduce the complex charged defect structures of zirconium oxide. To the best of our knowledge, the only experimental value of a defect formation energy for undoped tetragonal ZrO_2 is that of the $V''''_{Zr} + 2V^{\bullet\bullet}_{O}$ Schottky defect determined by Wang and Olander.[63] They performed a thermodynamic analysis for the thermogravimetric measurements of Xue[22] and obtained a value of 4.12 eV, while we obtained a DFT value of 5.53 eV for the same defect. Although such quantitative discrepancy is common between DFT and experimentally derived formation energies, it is worth mentioning that Wang and Olander relied in their analysis on the assumption that quadruply charged zirconium vacancies have to be the predominant native defect at high oxygen chemical potential to simplify the charge neutrality equation in this regime. The nature of the predominant defect in tetragonal ZrO_2 at high oxygen chemical potential is a long-standing controversy in the literature,[22,23] as discussed in the next subsection; hence, such an assumption is definitely questionable. Furthermore, in their analysis, Wang and Olander obtained negative entropy of formation for that particular defect, a result that they suspected in their concluding remarks. We will discuss more the potential origins of the discrepancy between DFT calculations and experiments in the next subsection.

B. Defect equilibria

In this subsection we discuss the equilibria of the electronic and point defects of tetragonal ZrO_2 utilizing Kröger–Vink diagrams. We further compare our results with the experimental results in the literature that attempt to explain the defect equilibria in tetragonal ZrO_2. We found that our results are consistent with experiments in terms of the predominant defect at low T ($\leqslant$1500 K) and low P_{O_2}($\leqslant 10^{-6}$ atm). In this regime, both experiments and our calculations confirmed the predominance of electrons followed by the doubly charged oxygen vacancies. Our computational results and experiments consistently predict the increase of electronic transference number by elevating the temperature. However, prior experiments could not resolve the type and the charge of the predominant defect at high P_{O_2}. Our DFT calculations were able to uncover the nature of this defect, and we found it to be the doubly charged oxygen vacancies at low temperatures and free electrons at high temperatures. Confidence in our conclusions is enhanced by the agreement between the slope of the predominant defect concentration in the calculated Kröger–Vink diagrams and the slope of conductance as a function of P_{O_2} in conductivity experiments.[23]

Figures 4(a) and 4(c) shows the calculated Kröger–Vink diagrams at 1500 and 2000 K, respectively, using the DFT band gap of 3.9 eV. Figure 4(b) is the calculated diagram at 1500 K using the experimental value[50] of 4.2 eV, and panel (d) is a reproduction of the conductivity measurements of Kofstad and Ruzicka.[23] In our calculated diagrams we show only the defects that have concentrations greater than 10^{-8} per ZrO_2 chemical formula. Moreover, we limit the horizontal axis to a range of P_{O_2} that extends from 1 to 10^{-15} atm, unless there is no value for the electron chemical potential in the band gap that achieves both charge neutrality and positivity of all formation energies. (See the Supplemental Material[32] for further details.) We chose those two representative temperatures for our calculated Kröger–Vink diagram because they represent two distinct classes (i.e., the low-temperature behavior is exemplified by the diagram at 1500 K and the high-temperature behavior is represented by the one at 2000 K). The distinction between the two behaviors is the predominant defect at each regime of P_{O_2} as detailed below. Before discussing the details of the diagrams, we introduce the notation log[D] for log of the defect concentration.

1. The Kröger–Vink diagram at 1500 K

The diagram at 1500 K using the DFT band gap in Fig. 4(a) can be divided into two regions. The first is a high P_{O_2} region that extends from 1 atm up to $10^{-1.4}$ atm, in which the doubly charged oxygen vacancies are the predominant defect. In this region charge compensation mainly takes place through the creation of quadruply charged zirconium vacancies and free electrons. The slope of log[D] for the predominant defect in this region is slightly negative but close to zero. This region is characteristic for ionic materials that predominantly form Schottky defects around the stoichiometric composition.[30] The second region extends from $10^{-1.4}$ atm up to 10^{-15} atm, and the predominant defects here are the free electrons followed by the doubly charged oxygen vacancies. Charge compensation mainly takes place among these two types of defects; hence, their log[D] have slopes of $(-1/6)$, as predicted by applying the law of mass action. This behavior of concentrations leads to an n-type electric conductivity. It is worth noting that the log[D] of neutral oxygen vacancies grows with a slope of approximately $(-1/2)$ by lowering P_{O_2} in this region until it becomes comparable to the concentration of free electrons. The slope of the log[D] of this defect is roughly independent of the P_{O_2} as it does not participate in achieving charge neutrality. The calculated diagram at 1500 K using the experimental value of 4.2 eV for the band gap is shown in Fig. 4(b). The same qualitative features described above still apply to this diagram. The major difference is that extending the band gap expands the horizontal region down to $10^{-5.7}$ atm. This is in better agreement with the conductivity measurements shown in Fig. 4(d) and discussed below. Increasing the band gap adds a penalty to the process of creating free electrons and holes. Hence the region in which free electrons predominates is delayed until a pressure as low as $10^{-5.7}$ atm is achieved.

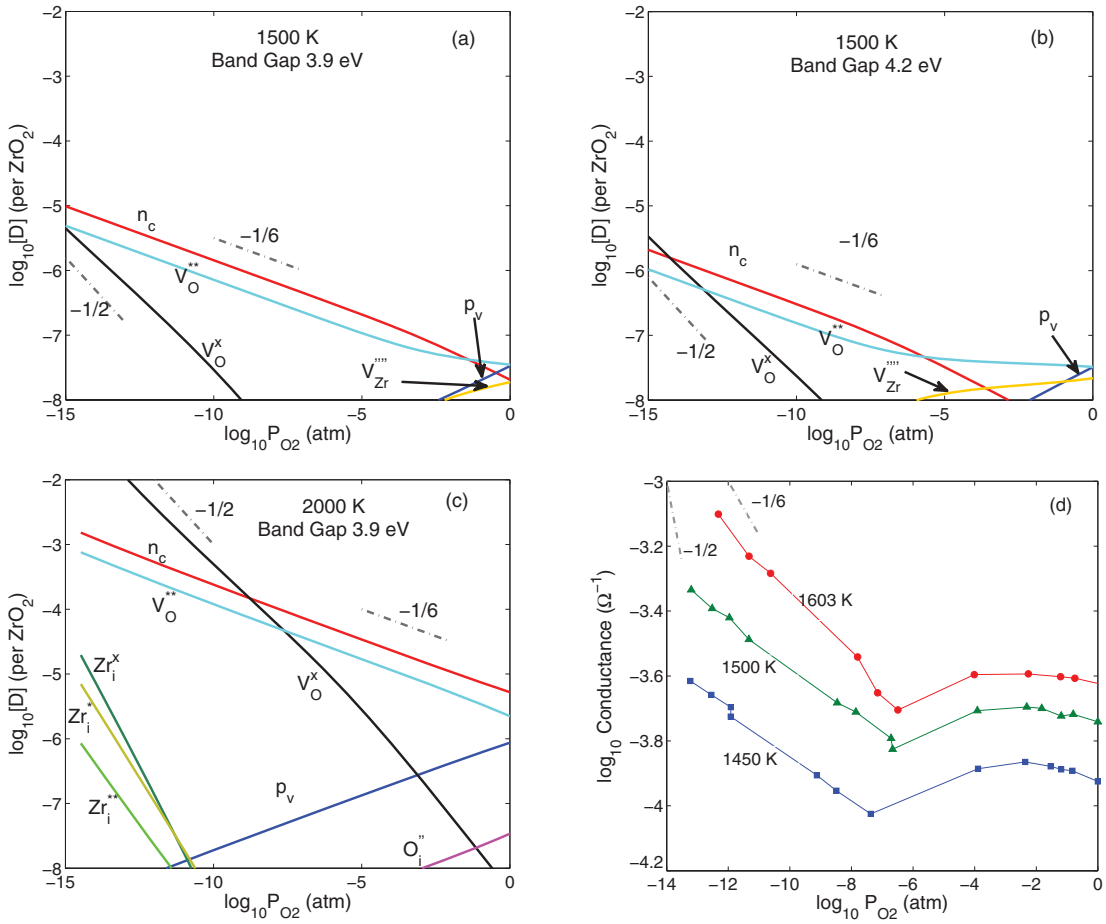

FIG. 4. (Color online) Calculated Kröger–Vink diagram for tetragonal ZrO_2 at (a) 1500 K using the DFT-predicted band gap of 3.9 eV, (b) 1500 K using the experimental band gap of 4.2 eV, (c) 2000 K using the DFT-predicted band gap. (d) Conductivity measurements on tetragonal zirconia by Kofstad and Ruzicka.[23] In panels (a–c) only the defects that have concentrations greater than 10^{-8} are shown. The dashed-dotted lines are guides for the eye showing the $(-1/2)$ and $(-1/6)$ slopes.

2. *The Kröger–Vink diagram at 2000 K*

Figure 4(c) shows the calculated Kröger–Vink diagram at 2000 K using the DFT band gap. The first distinction to note here compared with the lower temperature diagram is that more defects have concentrations greater than 10^{-8}, as is anticipated with a temperature-activated concentration. Second, in no region does $\log[D]$ of the predominant defect have a horizontal slope. This is an indication of an off-stoichiometric composition, and indeed it is hypostoichiometric, as we elaborate more in the next subsection. However, the diagram can still be divided into two regions. The first region extends from 1 atm to $10^{-8.8}$ atm. In this region the free electrons are predominant, followed by the doubly charged oxygen vacancies. Charge compensation takes place mainly among these two types of defects, hence, the $(-1/6)$ slope for both. In this region it is evident that doubly charged oxygen interstitials have higher concentrations compared with quadruply charged zirconium vacancies, indicating that if tetragonal ZrO_2 is stable at $P_{O_2} >$ 1 atm and can approach stoichiometric composition at 2000 K, then it would be an ionic compound in which the intrinsic region around the stoichiometric composition is dominated by oxygen Frenkel pairs. This is a significant difference compared with the finding at 1500 K, where the intrinsic region around the stoichiometric composition is dominated by Schottky defects. The second region in the Kröger–Vink diagram at 2000 K extends from $10^{-8.8}$ to roughly $10^{-14.5}$ atm. The border between the first and second region is due to the thermodynamic transition level for oxygen vacancies from charge state 2+ to 0, as shown in Fig. 3(d). As discussed above we did not find a value for the chemical potential of electrons that achieves both charge neutrality and positivity of the formation energies of all defects at pressures lower than $10^{-14.5}$ atm. In this region the neutral oxygen vacancies predominate, with a $\log[D]$ slope of approximately $(-1/2)$. Noticeably, the concentrations of three charge states of the zirconium interstitial grow in this region by lowering P_{O_2} but still below the concentrations of oxygen and electronic defects in this region.

3. *Comparison with experiments*

To compare our theoretical results with the experimental findings in the literature, we show in Fig. 4(d) the conductivity measurements of Kofstad and Ruzicka.[23] On an isotherm, the defect motilities are constant. Therefore, the log of the total conductance of a metal oxide exhibits a slope with P_{O_2} that is governed by the $\log[D]$ slope of the predominant defect. This

is what justifies the comparison between our DFT calculated concentrations and the conductivity measurements.

The conductivity measurements in Fig. 4(d) are in reasonable agreement with our calculated Kröger–Vink diagram at 1500 K in terms of exhibiting a slightly negative slope that is very close to zero at high P_{O_2}. At lower P_{O_2} the slope of the log of the conductance changes to roughly $(-1/6)$. It is important to note that according to these conductivity measurements, the first horizontal region extends down to 10^{-4} atm, then the $(-1/6)$ slope starts somewhere between 10^{-6} and 10^{-7} atm, depending on the temperature. This is more consistent with our calculated Kröger–Vink diagram using the experimental band gap of 4.2 eV.

The trends exhibited by the conductivity measurements shown here are also obeyed by the conductivity measurements of Vest and Tallan.[24] Additionally, thermogravimetric measurements by Xue[22] reproduced the $(-1/6)$ slope at low P_{O_2}. Thus, there is firm agreement that at low P_{O_2} the predominant defect is free electrons followed by doubly charged oxygen vacancies. However, the experimental efforts could not resolve the nature of the predominant defect at high P_{O_2}. Kofstad and Ruzicka, whose results are reproduced in Fig. 4(d), suggested that at high P_{O_2} both oxygen vacancies and interstitials are predominant. Vest and Tallan[24] could not deduce the predominant defect at high P_{O_2}. The thermogravimetric measurements of Xue[22] at high P_{O_2} exhibited a positive slope that can be fit well by either $+1/6$ (O_i'' predominant) or $+1/5$ (V_{Zr}'''' predominant) slopes, but the latter was chosen to conform to the classical potential predictions of Ref. 62. As we mentioned above and detailed in the Supplemental Material,[32] the results of this classical potential are highly nonphysical and cannot be used as a support of the nature of charged defects. We believe that the reasonable agreement between our DFT results and the conductivity measurements of Refs. 23 and 24 provide a solid ground to propose that this defect is the doubly charged oxygen vacancies at high P_{O_2}.

Our DFT calculations are consistent with the experiments in another aspect. We showed that at low temperatures (with respect to the range of stability of T-ZrO_2) electronic conductivity predominates in the region of low P_{O_2}. At higher temperatures, electronic conductivity predomination starts at high P_{O_2} and extends over a wide range of P_{O_2}. Measurements of the electronic and ionic transference numbers are consistent with this finding.[24,25]

4. Origins of the gap between theory and experiment

Beyond the finite temperature effects (which we demonstrated in Sec. II C and incorporated into our analysis), several other challenges obstruct improving the level of agreement between theory and experiments for determining the nature and concentrations of defects at a given thermodynamic state. On the experimental side, impurities are unavoidable and can significantly affect conductivity measurements. For example, the conductivity measurements of Guillot and Anthony[26] could not resolve the nature of the intrinsic predominant defects in T-ZrO_2 because of the interfering role of impurities. Moreover, to the best of our knowledge, all experiments on T-ZrO_2 to understand its defect equilibria were performed on polycrystalline samples. Grain boundaries and the associated space charge zones can significantly affect defect equilibria.[2] Not only the presence of grain boundaries, but also the size of the grains themselves can alter the equilibria.[64] The current DFT computational limits do not allow realistic calculations for charged defect energies at grain boundaries. Another factor that is pertinent to undoped T-ZrO_2 is that it is thermodynamically stable at very high temperatures, which poses limitations on the feasibility and accuracy of the measurements. On the theoretical side, the well-known DFT problem of underestimating the band gap has an exponentially amplified impact on determining the concentration of free electrons and holes, and hence on the overall equilibria of the charged defects. As shown in Figs. 4(a) and 4(b), applying a rigid shift of 0.3 eV to the conduction band improved agreement with the conductivity measurements. Applying rigid shifts to the conduction band to match the experimental band gap is a common practice.[65,66] However, the spread in the experimentally determined band gap values, as in the case of T-ZrO_2,[50,51] makes it difficult to choose one of these values with confidence. (See the Supplemental Material[32] for further discussion.) Furthermore, the well-known GGA-DFT issue of overbinding the oxygen molecule introduces an error of about 1.36 eV when using the PBE functional, as estimated by Wang *et al.*,[67] by fitting the formation enthalpy of simple nontransition metals to the experimental values. This overbinding impacts the accuracy of determining the chemical potential of oxygen. The outcome of applying such a correction is introducing a positive shift to the oxygen molecule energy. Hence, it becomes more favorable to create defects that lead to the incorporation of more oxygen gas into the solid. These defects are the oxygen interstitials and zirconium vacancies. This simultaneously reduces the concentration of oxygen vacancies and zirconium interstitials. In other words, the concentration of the negatively charged point defects increases and the concentration of the positively charged point defects decreases. This, in turn, reduces the concentration of free electrons and increases the concentration of free holes to maintain charge neutrality. Furthermore, the actual value of the correction depends strongly on the DFT simulation parameters, which restricts us to use of the simulation parameters of Ref. 67 in order to justify using 1.36 eV as a correction; otherwise, we have to repeat the fitting with our DFT parameters. Hence, we chose not to apply this correction to reduce the empiricism in our approach as much as possible.

C. Off-stoichiometry and electron chemical potential

The temperature and the oxygen partial pressure are the thermodynamic independent variables that determine the defect concentrations in the stress-free undoped T-ZrO_2, as discussed in the previous subsection. Two other important observables that are determined by T and P_{O_2} are the sample off-stoichiometry, x, and the chemical potential of electrons, μ_F. The former is an integral quantity that represents the collective defect equilibria and can be measured in thermogravimetric experiments.[22] The latter is a fundamental quantity that affects the transport and transfer of electrons and can be measured in electrochemical impedance spectroscopy experiments.[68] In our DFT calculations, the calculated defect concentrations can

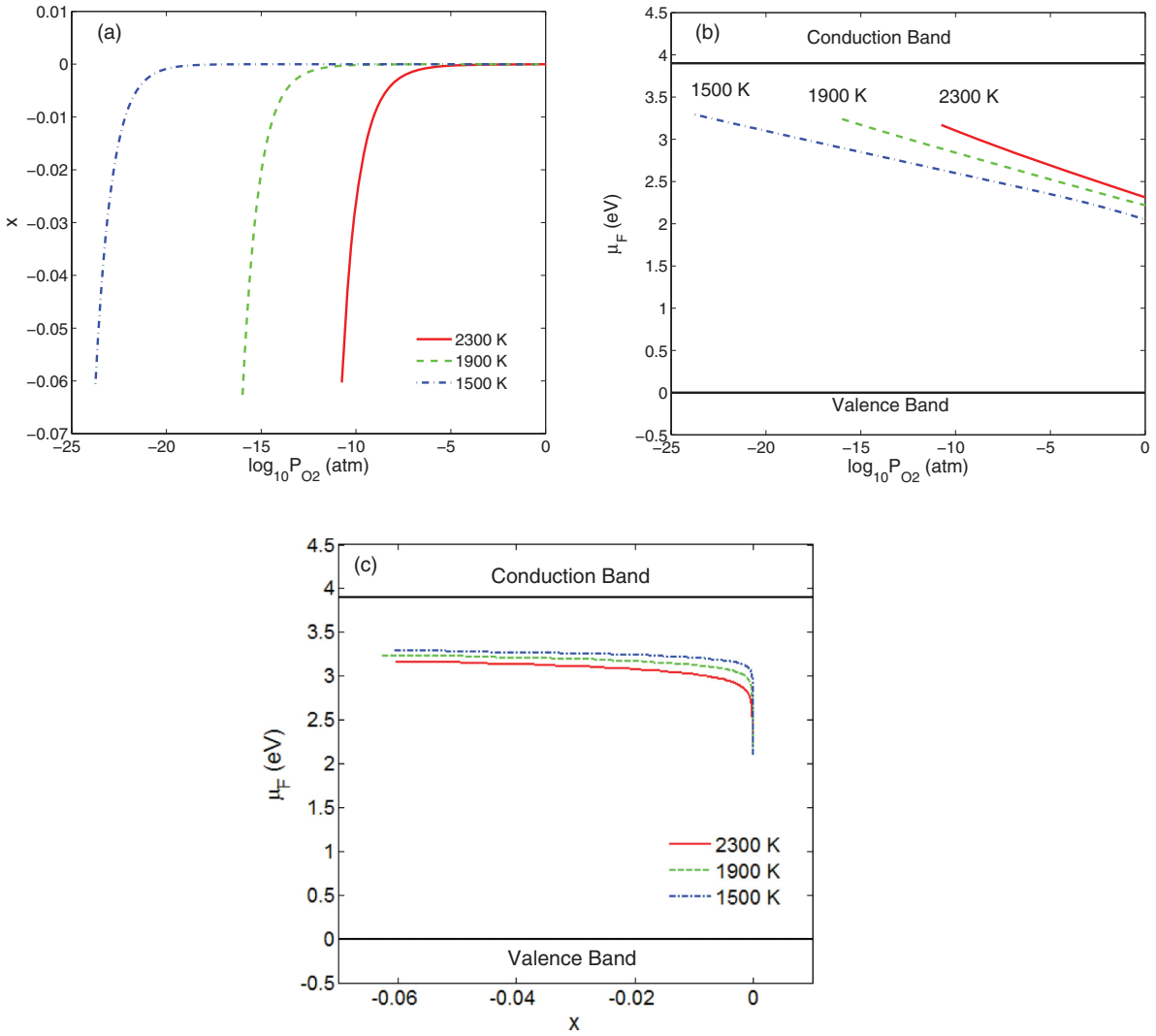

FIG. 5. (Color online) (a) Off-stoichiometry in ZrO_{2+x} as a function of the oxygen partial pressure, P_{O_2}, at different temperatures. The dependence of the electron chemical potential, μ_F, on the oxygen partial pressure at different temperatures (b), and on the off-stoichiometry, x, at different temperatures (c).

be used to compute x, and the charge neutrality condition determines the value of μ_F. In this subsection we discuss the relations among T-P_{O_2}-x-μ_F on the basis of our DFT calculations.

Figures 5(a)–5(c) depicts the relations among T-P_{O_2}-x-μ_F for stress-free undoped T-ZrO_2. Figure 5(a) shows that the off-stoichiometry of a T-ZrO_2 sample is almost independent of P_{O_2} for several orders of magnitude, and then at a very low P_{O_2}, the off-stoichiometry decreases significantly. The P_{O_2} at which $|x|$ (hypostoichiometry here) sharply increases is reduced with increasing the temperature.

As shown in the resulting Kröger–Vink diagrams, oxygen vacancies are responsible for the observed off-stoichiometry. This plot indicates also that T-ZrO_2 is always hypostoichiometric in the T-P_{O_2} range considered here, and this is consistent with the experimental results of Carnigila *et al.*[69] Figure 5(b) shows that, at 1 atm, the chemical potential of electrons μ_F is close to the middle of the DFT-calculated band gap. By lowering P_{O_2}, we find that μ_F moves toward the conduction band minimum within a range of about 1.3 eV. This is a self-doping effect due to the native defects, particularly the doubly charged oxygen vacancies, and is consistent with the n-type conductivity identified in the previous subsection. Finally in Fig. 5(c), we plotted the two dependent variables μ_F and x. Close to the stoichiometric composition, μ_F is very sensitive to very small changes in x. A very small change to a hypostoichiometric composition leads to a jump in the electron chemical potential toward the conduction band and then μ_F levels off by a further decrease of x. The value of x at which μ_F levels off in Fig. 4(c) corresponds to the value at which the sharp decrease in the off-stoichiometry takes place in panel (a). It is evident from panel (c) that the lower the temperature, the greater the maximum achievable electron chemical potential by self-doping.

The chemical potential of electrons in the bulk of a metal oxide controls the transport of electrons.[70] Furthermore, the transfer of electrons across interfaces is governed by the matching of μ_F across the interface. These two issues are of great importance in corrosion, catalysis, gate dielectrics, and dye-sensitized solar cells, among other applications. We

demonstrated here that our theoretical approach for defect equilibria in the bulk of a metal oxide consistently determines the value of μ_F according to the surrounding thermodynamic conditions. While we did not address here determining the variations of μ_F across an interface, we believe that our determination for μ_F in bulk zirconia sets a necessary boundary condition needed for the accurate determination of its variations at interfaces with the same approach described here.

D. Defect atomic and electronic structures

So far our discussion of the point defects was based on energetics. However, we believe it is also important to elucidate the atomic and electronic structure of the defect because the atomic and electronic characteristics of each defect provide a signature to help in detecting it experimentally. Moreover, these characteristics are needed to understand the transport kinetics of the point defects and their effect on the mechanical properties of the material and, thus, can be important for future work concerned with these defects. In this subsection we describe our key observations related to the atomic and the electronic structures of the major defects that we assessed in the analysis presented above, starting with vacancies and ending with interstitials. More details are presented in the Supplemental Material.[32]

Several experiments and DFT calculations confirmed that oxygen vacancies in all phases of zirconium oxide have the ability to trap electrons forming F-centers.[27,28,65,71] Our calculations are consistent with these prior reports, as we observed electron localization on the vacant oxygen site for both V_O^x and $V_O^\bullet$. The ground state for the two electrons localized in the V_O^x defect is singlet. It is energetically very unfavorable for zirconium cations to be reduced to the oxidation state 3+; hence, the electrons get trapped in the vacant site. On the other hand, we observed in our calculations that all zirconium vacancies except $V_{Zr}^{''''}$ lead to the formation of the antimorph of the F-center, namely the V-center. In a V-center, a hole is trapped in an oxide ion that is the nearest neighbor of a cation vacancy, which means each of the zirconium vacancies other than $V_{Zr}^{''''}$ is a cluster of point defects. For example, a more accurate notation for V_{Zr}^x would be $(4O_O^\bullet V_{Zr}^{''''})^x$; however, we kept the former as a simplified notation throughout the paper. As we showed in the previous subsections, the V-centers associated with all zirconium vacancies except $V_{Zr}^{''''}$ always have a minute concentration.

As indicated in the computational methods section, we considered seven different structures of the interstitial oxygen at all charge states after excluding the ⟨100⟩ crowdion. Our calculations indicate that for the charge states 0 and −1, the ⟨110⟩ split dumbbell is the energetically most favorable structure. The lengths of the dumbbells are 1.47 and 1.99 Å, respectively, larger than our calculated bond length for the oxygen molecule, which is 1.23 Å. Examination of the charge density (see the Supplemental Material[32]) shows that the charge is almost evenly distributed on the two ions of the dumbbell. Moreover, for the charge state 0, a very distinct feature appears in the phonon DOS (see the Supplemental Material[32]), which is a characteristic peak at a frequency of 939 cm^{-1}. Although this peak is at a frequency much higher than any vibrational mode in the crystal, it is still much lower than 1560 cm^{-1}, our calculated vibrational frequency for the oxygen molecule. On the other hand, for the charge state 2−, the octahedral site is the most favorable. The strong columbic repulsion does not allow the doubly charged interstitial oxygen ion to have a dumbbell or a crowdion configuration.

IV. CONCLUSION

In this paper we presented a framework to understand and predict the equilibria of charged point defects in a metal oxide, exemplified here by the technologically important case of T-ZrO_2. In our treatment we considered the relevant finite temperature excitations due to phonon vibrations and electronic entropy and demonstrated these to be important for the quantitative accuracy of the DFT calculations. Phonons add a significant contribution to the free energy of formation of the defects, especially at high temperatures. The electronic entropy contribution gets enhanced when the chemical potential of electrons (Fermi level) coincides with a defect state in the band gap.

We found out that T-ZrO_2 is hypostoichiometric over its range of stability. Within the temperature range (1440–2640 K) of stability of T-ZrO_2, our calculations using the DFT band gap of 3.9 eV (*experimental band gap of 4.2 eV*) indicated that: (a) at low temperatures up to 1500 K (*up to 1700K*) the doubly charged oxygen vacancies predominates at high P_{O_2} and electrons predominates at low P_{O_2}, (b) at temperatures higher than 1500 K (*higher than 1700K*) the electrons are the dominant defects at high P_{O_2} within a wide range, while the neutral oxygen vacancies (F-centers) predominate at low P_{O_2}. Our computational predictions are consistent with prior experiments that indicated the predominance of electrons at low temperatures and low P_{O_2} and confirmed the increase of the electronic transference number at higher temperatures. However, the predominant point defect type and charge state were not previously amenable to experimental determination; hence, our computational results here filled a gap in our understanding of zirconia defect equilibria over a wide range of thermodynamic conditions.

The chemical potential of electrons, which is important for electron transport and transfer, was determined with the charge neutrality condition. For the undoped T-ZrO_2 studied here, μ_F, was found to reside close to the middle of the band gap at high P_{O_2} and to approach the conduction band edge by lowering the P_{O_2} due to a self-doping effect by oxygen vacancy defects.

It is possible to extend the thermodynamic space from (T, P_{O_2}) to include mechanical strain, doping, and electric and magnetic fields. By constructing diagrams similar to the traditional Kröger–Vink diagram, one can optimize the thermodynamic conditions that extremize certain desirable or undesirable point defect types for technological purposes, such as catalysis, corrosion, and microelectronics applications. An accurate determination of these optimum conditions dictates careful sampling for the excitations and degrees of freedom of the system under consideration.

ACKNOWLEDGMENTS

This research was supported by the Consortium for Advanced Simulation of Light Water Reactors, an Energy Innovation Hub for Modeling and Simulation of Nuclear

Reactors, under US Department of Energy Contract No. DE-AC05-00OR22725. We acknowledge the National Science Foundation for computational support through the TeraGrid Advanced Support Program with research allocation TG-DMR110004 and the XSEDE Science Gateways program with research allocation TG-DMR120025.

*Corresponding author: byildiz@mit.edu

[1] J. Frenkel, Z. Phys. **35**, 625 (1926).
[2] H. Tuller and S. Bishop, Annu. Rev. Mater. Res. **41**, 269 (2011).
[3] A. Cheroneos, B. Yildiz, A. Tarancon, D. Parfitt, and J. Kilner, Energy and Environ. Sci. **4**, 2774 (2011).
[4] D. Marrocchelli and B. Yildiz, J. Phys. Chem. C **116**, 2411 (2012).
[5] A. L. Shlugerand and Adam Foster, *in Nano and Giga Challenges in Microelectronics*, edited by J. Greer, A. Korkin, and J. Labanowski (Elsevier, Amsterdam, 2003).
[6] D. R. Olander, *Fundamental Aspects of Nuclear Reactor Fuel Elements* (Technical Information Center, Energy Research and Development and Administration, Spingfield, VA, 1976).
[7] M. Youssef and B. Yildiz, unpublished (2012).
[8] J. Chevalier, L. Gremillard, A. V. Virkar, and D. R. Clarke, J. Am. Ceram. Soc. **92**, 1901 (2009).
[9] W. C. Maskell, Solid State Ionics **134**, 43 (2000).
[10] U. Schulz, C. Leyens, K. Fritscher, M. Peters, B. Saruhan-brings, O. Lavigne, J. M. Dorvaux, M. Poulain, R. Merval, and M. Caliez, Aeroscpace Sci. Technol. **7**, 73 (2003).
[11] A. Yilmazbayhan, E. Breval, A. T. Motta, and R. J. Comstock, J. Nuc. Mat. **349**, 265 (2006).
[12] D. D. Macdonald and G. Engelhardt, ECS Transactions Transactions **28**, 123 (2010).
[13] Y. Chen, M. Urquidi-Macdonald, and D. D. Macdonald, J. Nucl. Mater. **348**, 133 (2006).
[14] B. Cales, Clin. Orthopaed. Relat. Res. **379**, 94 (2000).
[15] I. Denry and J. R. Kelly, Dent. Mater. **24**, 299 (2008).
[16] X. Guo, Chem. Mater. **16**, 3988 (2004).
[17] S. Hull, Rep. Prog. Phys. **67**, 1233 (2004).
[18] A. Kushima and B. Yildiz, J. Mater. Chem. **20**, 4809 (2010).
[19] D. J. Green, R. H. Hannik, and M. V. Swanik, *Transformation Toughening of Ceramics* (CRC Press, Boca Raton, FL, 1989).
[20] X. Zhao and D. Vanderbilt, Phys. Rev. B **65**, 075105 (2002).
[21] D. D. Macdonald, Pure. Appl. Chem. **71**, 951 (1999).
[22] J. Xue, J. Electrochem. Soc. **138**, 36C (1991).
[23] P. Kofstad and D. J. Ruzicka, J. Electrochem. Soc. **110**, 181 (1963).
[24] R. W. Vest and N. M. Tallan, J. Am. Ceram. Soc. **48**, 472 (1965).
[25] L. A. McClaine and C. P. Coppel, J. Electrochem. Soc. **113**, 80 (1966).
[26] A. Guillot and M. Anthony, J. Sol. Stat. Chem. **15**, 89 (1975).
[27] A. Eichler, Phys. Rev. B **64**, 174103 (2001).
[28] M. Ganduglia-Pirovano, A. Hofmann, and J. Sauer, Surf. Sci. Rep. **62**, 219 (2007).
[29] A. A. Safonov, A. A. Bagatur'yants, and A. A. Korkin, Microelectron. Eng. **69**, 629 (2003).
[30] Y.-M. Chiang, D. P. Birnie III, and W. D. Kingery, *Physical Ceramics: Principles for Ceramic Science and Engineering* (John Wiley & Sons, New York, 1997).
[31] G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).
[32] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.86.144109 for further computational details and additional results and notes.
[33] D. S. Sholl and J. A. Steckel, *Density Functional Theory: A Practical Introduction* (Wiley & Sons, Hoboken, NJ, 2009).
[34] M. W. Chase, Jr., *NIST-JANAF Thermochemical Tables* (American Institute of Physics, Woodbury, NY, 1998).
[35] J.-P. Crocombette, D. Torumba, and A. Chartier, Phys. Rev. B **83**, 184107 (2011).
[36] X. Li, M. W. Finnis, J. He, R. K. Behera, S. R. Phillpot, S. B. Sinnot, and E. C. Dickey, Acta Mater. **57**, 5882 (2009).
[37] S. Kasamatsu, T. Tada, and S. Watanabe, Solid State Ionics **183**, 20 (2011).
[38] N. Aschcroft and N. Mermin, *Solid State Physics* (Brooks/Cole, Belmont, CA, 1976).
[39] F. Zhou, T. Maxisch, and G. Ceder, Phys. Rev. Lett. **97**, 155704 (2006).
[40] B. T. Webber, M. C. Per, D. W. Drumm, L. C. L. Hollenberg, and S. P. Russo, Phys. Rev. B **85**, 014102 (2012).
[41] C. Wolverton and A. Zunger, Phys. Rev. B **52**, 8813 (1995).
[42] R. Pässler, J. App. Phys. **89**, 6235 (2001).
[43] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[44] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).
[45] G. Kresse and J. Hafner, Phys. Rev. B **49**, 14251 (1994).
[46] G. Kresse and J. Furthmüller, Comput. Mater. Sci **6**, 15 (1996).
[47] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
[48] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[49] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396 (1997).
[50] D. W. McComb, Phys. Rev. B **54**, 7094 (1996).
[51] R. H. French, S. J. Glass, F. S. Ohuchi, Y. N. Xu, and W.Y. Ching, Phys. Rev. B **49**, 5133 (1994).
[52] V. I. Anisimov and F. Aryasetiawan, J. Phys. Condens. Matter **9**, 767 (1997).
[53] R. Ramprasad, H. Zhu, P. Rinke, and M. Scheffler, Phys. Rev. Lett. **108**, 066404 (2012).
[54] A. P. Scott and L. Random, J. Phys. Chem. **100**, 16502 (1996).
[55] J. X. Zheng, G. Ceder, T. Maxisch, W. K. Chim, and W. K. Choi, Phys. Rev. B **75**, 104112 (2007).
[56] K. Momma and F. Izumi, J. Appl. Crystallogr. **44**, 1272 (2011).
[57] A. Togo, F. Oba, and I. Tanaka, Phys. Rev. B **78**, 134106 (2008).
[58] P. E. Blöchl, O. Jepsen, and O. K. Andersen, Phys. Rev. B **49**, 16223 (1994).
[59] S. F. J. Cox, J. L. Gavartin, J. S. Lord, S. P. Cottrel, J. M. Gil, H. V. Alberto, J. P. Duarte, R. C. Vilao, N. A. de Campos, D. J. Keeble, E. A. Davis, M. Charlton, and D. P. van der Werf, J. Phys.: Condens. Matter **18**, 1079 (2006).
[60] G. Mahan, Solid State Ionics **1**, 29 (1980).
[61] C. Van de Walle and J. Neugebauer, J. Appl. Phys. **95**, 3851 (2004).
[62] A. Dwivedi and A. N. Cormack, Philos. Mag. A **61**, 1 (1990).
[63] W. Wang and D. Olander, J. Am. Ceram. Soc. **76**, 1242 (1993).

[64]J. L. M. Rupp, Solid State Ionics **207**, 1 (2012).
[65]A. S. Foster, V. B. Sulimov, F. Lopez Gejo, A. L. Shluger, and R. M. Nieminen, Phys. Rev. B **64**, 224108 (2001).
[66]D. Shrader, S. M. Khalil, T. Gerczak, T. R. Allen, A. J. Heim, I. Szlufarska, and D. Morgan, J. Nuc. Mat. **408**, 257 (2011).
[67]L. Wang, T. Maxisch, and G. Ceder, Phys. Rev. B **73**, 195107 (2006).
[68]P. S. Archana, R. Jose, M. M. Yusoff, and S. Ramakrishna, App. Phys. Lett. **98**, 152106 (2011).
[69]A. C. Carniglia, S. D. Brown, and T. F. Shroeder, J. Am. Ceram. Soc. **54**, 13 (1971).
[70]J. Bisquert, J. Phys. Chem. C **111**, 17163 (2007).
[71]D. Simeone, G. Baldinozzi, D. Gosset, S. LeCaër, and L. Mazerolles, Phys. Rev. B **70**, 134116 (2004).