First-principles study of the valence band offset between silicon and hafnia

Blair R. Tuttle*
Department of Physics, Penn State Erie—The Behrend College, Erie, Pennsylvania 16563, USA

Chunguang Tang and R. Ramprasad
Department of Chemical, Materials and Biomolecular Engineering, Institute of Materials Science, University of Connecticut, 97 N. Eagleville Road, Storrs, Connecticut 06269 USA

(Received 28 September 2006; revised manuscript received 28 March 2007; published 22 June 2007)

First-principles density-functional calculations are used to examine the interface structures for crystalline hafnia strained to fit epitaxially on Si(001). The valence band offset has been calculated for several model heterojunctions. The results are compared to experiments and previous calculations. Interface structure and HfO_2 band tails are found to be important for band offset formation.

DOI: 10.1103/PhysRevB.75.235324 PACS number(s): 71.15.Nc, 74.25.Jb, 79.60.Jv

I. INTRODUCTION

Interfaces between silicon and oxides are of great fundamental and technological interest. According to the International Technology Roadmap for Semiconductors, the SiO_2-based gate oxide in metal-oxide-semiconductor (MOS) transistors used in computer processors should be replaced by a higher-K value dielectric material starting in 2008.[1] Presently, off-state gate tunneling is increasing power consumption and degrading performance. Many obstacles exist for incorporating alternative gate dielectrics into MOS transistors. Finding dielectric materials with sufficiently high band offsets to limit gate tunneling is essential. Metal oxides, such as zirconia (ZrO_2), hafnia (HfO_2), and others, have generated interest. Recently, hafnia has emerged as the leading candidate for a replacement gate dielectric.[2,3] In addition to having dielectric constants approximately five times SiO_2, hafnia-based oxides are structurally stable on silicon during fabrication and operation. This interest in hafnia for gate dielectric applications has motivated a desire to understand the fundamental properties of hafnia-based bulk and interfacial systems.

Several groups have previously investigated the bulk properties of hafnia. Bulk physical, electronic, and dielectric properties have been studied with first-principles density-functional theory[4–7] and are well understood. While the Si/HfO_2 heterojunctions have been investigated experimentally, they are not well understood. Several groups have completed theoretical studies of silicon-zirconia interfaces;[8–11] however, less attention has been given to interfaces between silicon and hafnia.[8,9] Moreover, in the two studies reporting valence band offsets, there is a 2 eV spread in the values reported. Overall, the physics controlling the silicon-hafnia heterojunction band diagram is poorly understood.

In this paper, we present density-functional calculations for interfaces of silicon- and hafnia-based oxides. We construct models of Si(001) slabs with varying layers of strained phases of hafnia on top. Two types of oxygen passivation are considered. The goal of the present study is to help elucidate the fundamental physics relevant to the $Si(001)/HfO_2$ interface and its band diagram.

II. MODELS AND METHODOLOGY

For our calculations, we use first-principles density-functional methods as implemented in the VASP code.[12,13] For all calculations, we use a plane-wave basis, with a 350 eV cutoff energy. For the exchange-correlation potential, we employ the generalized gradient approximation (GGA) of Perdew and Wang.[14] We have performed bulk and interface calculations using supercells of varying size. Integrations over the Brillioun zone were performed over a grid of special k points. In all cases, k-point sampling was increased until the convergence of the total energy was approximately 0.01 eV; e.g., for an interface supercell with dimensions of 5.46

TABLE I. Properties of various polymorphs of HfO_2. Results in brackets are from previous studies.

Structure	Lattice vectors (Å)	$\Delta E/HfO_2$	Band gap (eV)	ΔVBM (eV)
Monoclinic	[5.13, 5.19, 5.30][a]		3.82 [3.80][a]	3.62
Tetragonal	[3.58, 5.20][a]	0.16 [0.16][a]	4.50 [4.45][a]	3.41
Cubic	[5.13][a]	0.25 [0.24][a] [0.25][b]	3.65 [3.70][a]	2.45
Epi-cubic	5.46, 4.92	0.95	2.09	2.52
Epi-tetra	5.46, 5.16	0.85	3.00	3.93
Epi-mono	5.46, 4.90	0.18 [0.16][b]	4.41	4.35

[a]Reference 6.
[b]Reference 8.

$\times 5.46 \times 50$ Å^3, a $4 \times 4 \times 1$ k-point sampling was used. The theoretical lattice constant of silicon ($a=5.46$ Å) was used throughout. Nuclei were represented by ultrasoft pseudopotentials where the cutoff radii are 1.80, 3.05, and 1.40 a.u. for silicon, hafnium, and oxygen, respectively.[13,15] All positions were allowed to relax until a force tolerance of 0.05 eV/Å on each atom was reached. For the valence band offsets reported here, we use the standard bulk plus lineup method which involves comparing bulk eigenvalues corrected by a dipole term calculated for a heterojunction model.[16] For the valence band offset (VBO) values reported, the numerical uncertainty is less than 0.1 eV. However, fundamental uncertainties can be larger as discussed below.

We explored common and hypothetical polymorphs whose properties are reported in Table I. To construct our interface models, we began with the structures reported by Jaffe *et al.*[6] for the monoclinic, tetragonal, and cubic phases. Very slight changes occurred when we performed independent relaxations of the respective structures. As illustrated in Table I, the agreement with previous work[6,8] is excellent; we found the monoclinic phase to be the lowest in energy, whereas the tetragonal and cubic phases are higher in energy by 0.16 and 0.25 eV/HfO_2, respectively. We explored the minimum electronic band gap within GGA by examining the eigenvalues along high-symmetry directions and over a dense grid. The gap for each polymorph considered is reported in column four of Table I. The minimum band gap of HfO_2 covers a wide range of values from 2.0 to 4.5 eV. The gap values for the stable polymorphs are all within 0.05 eV of previously reported values.[6] For bulk plus lineup band offset calculations,[16] one needs to know the difference between the bulk silicon and the oxide valence band maxima (ΔVBM). These values are reported in the fifth column of Table I. For all polymorphs, the HfO_2 valence band maximum is below the value for bulk silicon. As with the bandgap values, the ΔVBM values cover a wide range, in this case from 2.5 to 4.4 eV. The variation in the electronic properties of HfO_2 polymorphs is not unexpected since the oxygen bonding and coordination change dramatically between the polymorphs. This issue is discussed below in the context of epitaxial oxides.

Bulk strained oxides were created to be placed epitaxially on top of the Si(001) surface. For each bulk stable polymorph, two lattice vectors are expanded to fit onto silicon (001). Allowing the third lattice vector and all coordinates to fully relax, a bulk epitaxial oxide was created. Only the lowest-energy strained polymorph for each stable phase was considered. For the monoclinic structure, we also rectified the lattice vectors, thereby making them orthogonal. Consistent with earlier studies, we find that the effect of rectification is negligible.[8] The final structures are shown in Fig. 1 including (a) the epitaxial cubic (epi-cubic), (b) the epitaxial tetragonal (epi-tetra), and (c) the epitaxial monoclinic (epi-mono) phases. The final (001) lattice vectors are 4.92, 5.16, and 4.90 Å for the epi-mono, epi-tetra, and epi-cubic bulk structures, respectively. Compared to the energy per HfO_2 unit of the monoclinic phase, these epitaxial models are 0.18, 0.85, and 0.94 eV higher for the epi-mono, epi-tetra, and epi-cubic models, respectively. Interestingly, the epi-mono polymorph has a modest strain energy compared to the other two strained polymorphs considered.

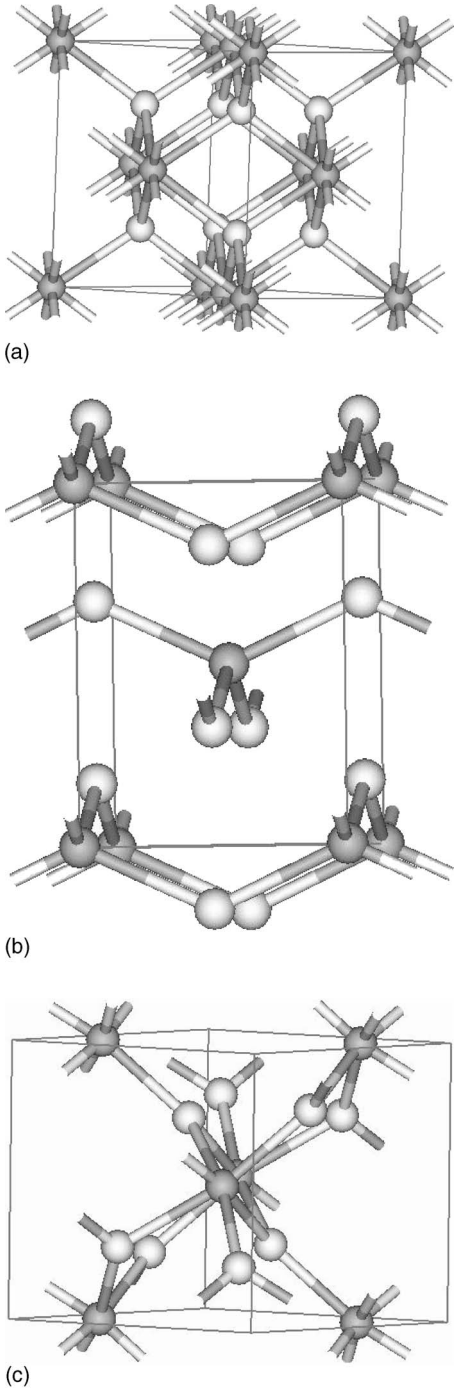

FIG. 1. Ball-and-stick models of (a) epi-cubic, (b) epi-tetra, and (c) epi-mono phases of hafnia, all strained to fit epitaxially on Si(100) (Hf=gray and O=white).

III. RESULTS AND ANALYSIS

Clearly, thermodynamic considerations strongly favor the epi-mono polymorph over the other two phases for epitaxial formation. Nevertheless, we constructed interface models from each epi-polymorph varying the number of HfO_2 layers between 1 and 8. Following previous research,[8,9] we first considered heterojunction models with no vacuum. Since the dimer rows are perpendicular to each other at each interface,

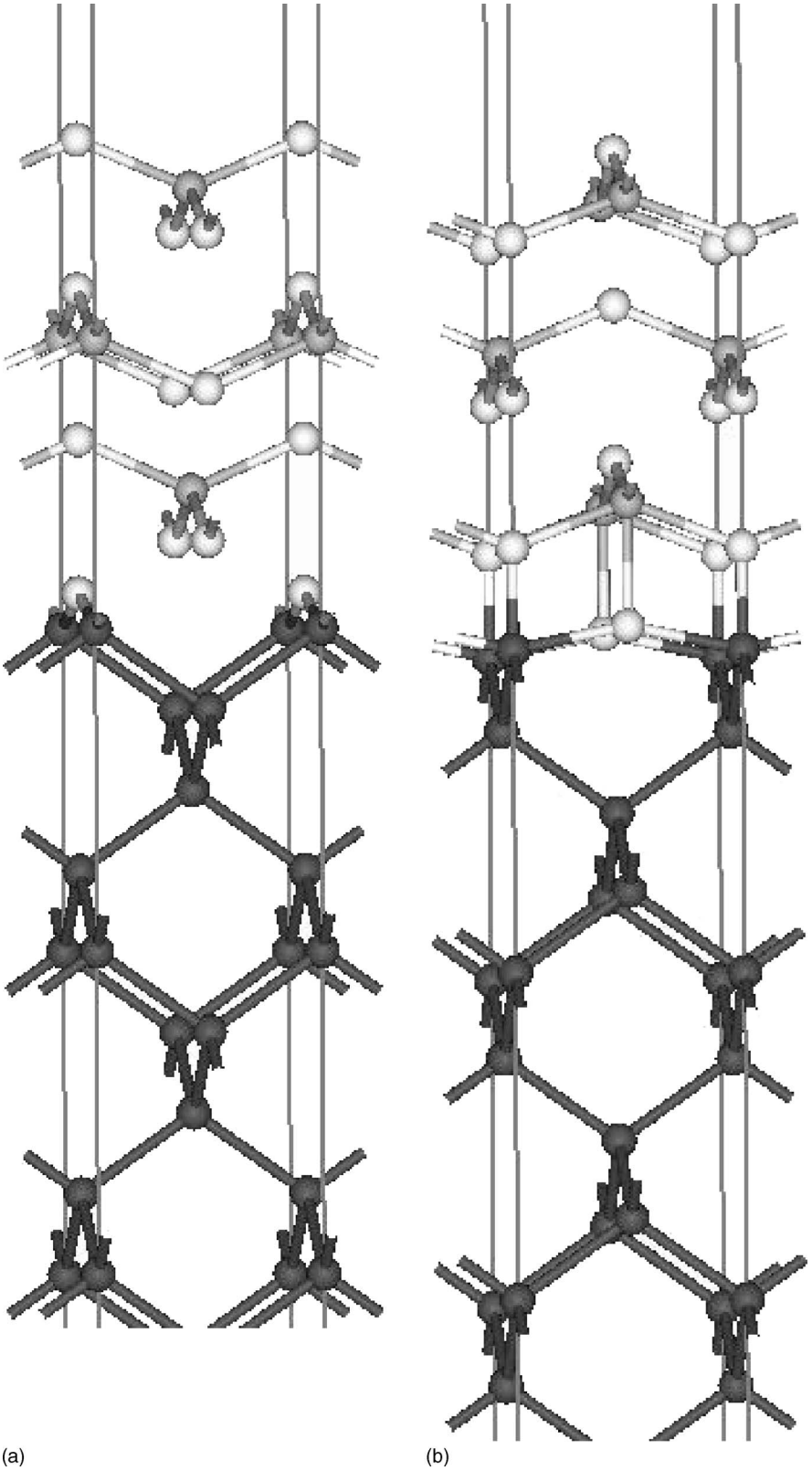

FIG. 2. Ball-and-stick models of (a) the O4 interface and (b) the O3 interface of epi-tetra HfO_2 on top of Si(100) (Hf=gray, O=white, and Si=dark).

such models involved asymmetric relaxations from the interface into the bulk. This prevented an epi-cubic or epi-tetra bulk phase from forming. To avoid this problem, we constructed Si/HfO_2/vacuum interface models. The interface between HfO_2 and vacuum does not impose contrary relaxations. The epi-tetra slab readily forms a bulklike region in the middle of the slab. We also considered interface models with symmetric HfO_2 layers on both surfaces of the silicon slab. These models were created to ensure accurate valence band offset calculations without any interfering electric fields. We found passivating one surface with oxygen only, without any Hf atoms, did lead to small electric fields due to

the differing interfacial dipoles. While various interfacial bonding arrangements are possible, we focus on the two interfaces previously found to be the lowest in energy for most values of the oxygen chemical potential.[9,11]

First, consider the epi-cubic polymorph. All oxygens are fourfold coordinated with the Hf-O bond lengths equal to 2.31 Å, slightly longer than the values for the stable phases of HfO_2. This polymorph fits on Si(001) with a minimum (1×1) surface unit cell with sides of length $a/\sqrt{2}$. While the epi-cubic polymorph was the basis of a study by Peacock *et al.*,[9] we found that it was not stable on Si(001). The epi-cubic hafnia transformed into the epi-tetra polymorph upon relaxation. Dong *et al.*[11] also found that an epi-tetra version of zirconia was stable on Si(001). Since the epi-cubic phase was not stable on Si(001), a valence band offset could not be reliably derived.

Next, consider the epi-tetra polymorph which also fits on the minimum (1×1) surface unit cell with sides of length $a/\sqrt{2}$. For the stable tetragonal phase, all oxygens are fourfold coordinated and all hafnia are eightfold coordinated. However, half the Hf-O bonds are short at 2.07 Å and half are long at 2.38 Å. In the epi-tetra phase, the system splits further with half the oxygen atoms moving up and the other half moving down, so that the material now forms sheets of HfO_2 units, as illustrated in Fig. 1(b). For the epi-tetra phase, there are still four bonds, but two bonds are strong with lengths of 2.15 Å and two are nominal with bonds 2.50 Å long.

Following the work of Peacock *et al.*,[9] we construct interface models where the valence shells of all atoms are satisfied by an oxygen interfacial passivation (Si-Si-O O-Hf-O O-Hf-O…). In Fig. 2(a), we show a HfO_2 constructed with the Hf above an open interstitial in the silicon surface. This is called the O4 interface by Peacock *et al.*[9] since the interfacial oxygen atoms are nominally fourfold coordinated. The structure in Fig. 2(b) includes a Hf atom directly above a bridging oxygen atom. This interface is called O3 since the interfacial oxygen atoms are threefold coordinated. We focused on these interfaces since they were the lowest in energy for oxygen rich conditions typical of growth.[9,11] We searched but were unable to find any lower-energy oxygen rich interfaces.

We constructed interface models in a manner similar to epitaxial growth, i.e., one HfO_2 layer at a time. All models included 17 layers of silicon and between one and eight layers of epi-tetra HfO_2. All atoms are allowed to relax; however, for both heterojunctions, only atoms in the few layers around the interface diverge from the respective crystalline coordinates. Compared to the O3 interface, we find the O4 interface model to be 2.0 eV higher in energy per $Si(1\times1)$ unit for the largest models. This result contrasts with that of Peacock *et al.*,[9] who find the O4 interface to be only 0.2 eV higher. The discrepancy here may be due to the fact that Peacock *et al.*[9] appear to force their HfO_2 slab portion to be epi-cubic. To test our result, we examined similar models substituting Zr for Hf since more work has been done on ZrO_2 interfaces. Our $Si(100)/ZrO_2$/vacuum results show the O4 interface to be 1.8 eV higher than the O3 interface. This is close to the result of Dong *et al.*[11] (1.6 eV) but much higher than that of Peacock *et al.*[9] (0.4 eV).

TABLE II. DFT valence band offset results in eV for two polymorphs of HfO_2 on Si(001). Results for two interface terminations (O3 and O4) are reported.

VBO (eV)	O3	O4
Epi-tetra	3.2	2.0
Epi-mono		3.1

Epitaxial growth will be sensitive to the energetics of the first few atomic layers. We find the epi-tetra O3 interface to be still about 2 eV lower in energy than the O4 interface even down to one HfO_2 layer. These results suggest the O3 interface should be favored for epitaxial growth. Below, we will revisit this issue in the context of our epi-mono interface calculations.

Using the standard bulk plus lineup method,[16] we determine the valence band offsets for the epi-tetra models with eight HfO_2 layers. The VBO is found to be 2.0 eV for the O4 interface and 3.2 eV for the O3 interface, as reported in Table II. Our results show that the dipole is large and that the placement of interfacial oxygens significantly affects the interfacial dipole. In contrast, Peacock *et al.*[9] find only a mild difference in the VBO results, with a VBO of 2.4 eV for the O4 interface and 2.3 eV for the O3 interface. The study of Dong *et al.*,[11] however, finds a large variation in the O3 and O4 VBOs for $Si(001)/ZrO_2$ interfaces, consistent with the present findings for the closely related $Si(001)/HfO_2$ interfaces.

Now, consider the epi-mono HfO_2 phase [Fig. 1(c)] and the interfaces it makes with silicon. The bulk epi-mono structure is similar to the monoclinic phase. However, the oxygen atoms are now all threefold coordinated with the Hf-O bond lengths between 2.05 and 2.15 Å. The O4 and O3 interface models are shown in Figs. 3(a) and 3(b), respectively. Each model heterojunction has a 17 Si layer slab, with six HfO_2 layers on top and a simple O-H passivation of the bottom. Only the top portions of the heterojunctions are shown in Fig. 3. Unlike in the case of the epi-tetra model (Fig. 2), here the O3 interface causes a transformation of the bulk HfO_2 epi-mono structure. The phase transformation causes the HfO_2/vacuum interface in the O3 model to differ dramatically from the other models. Only in the epi-mono O3 heterojunction does the vacuum surface oxygen atom break from being twofold coordinated; instead, the surface oxygen points straight up and is singly coordinated. The O4 epi-mono interface is 1.5 eV/(Si 1×1 unit) more favorable than the respective O3 interface. Examining interfaces with four HfO_2 layers, we find the O4 interface to be still more favored by about 1.7 eV/$(1\times1$ unit). This indicates that the $Si:HfO_2$ interface bonding is the main source of the discrepancy between the O4 and the O3 interface energetics. Equilibrium thermodynamics therefore favors the formation of the epi-mono O4 interface under oxygen rich conditions, consistent with experiments. Of course, reaction kinetics can favor a higher-energy interface as discussed by Peacock *et al.*[9]

The relaxations of the O3 interface generate distinct dipoles at the three interfaces and large electric fields ensue.

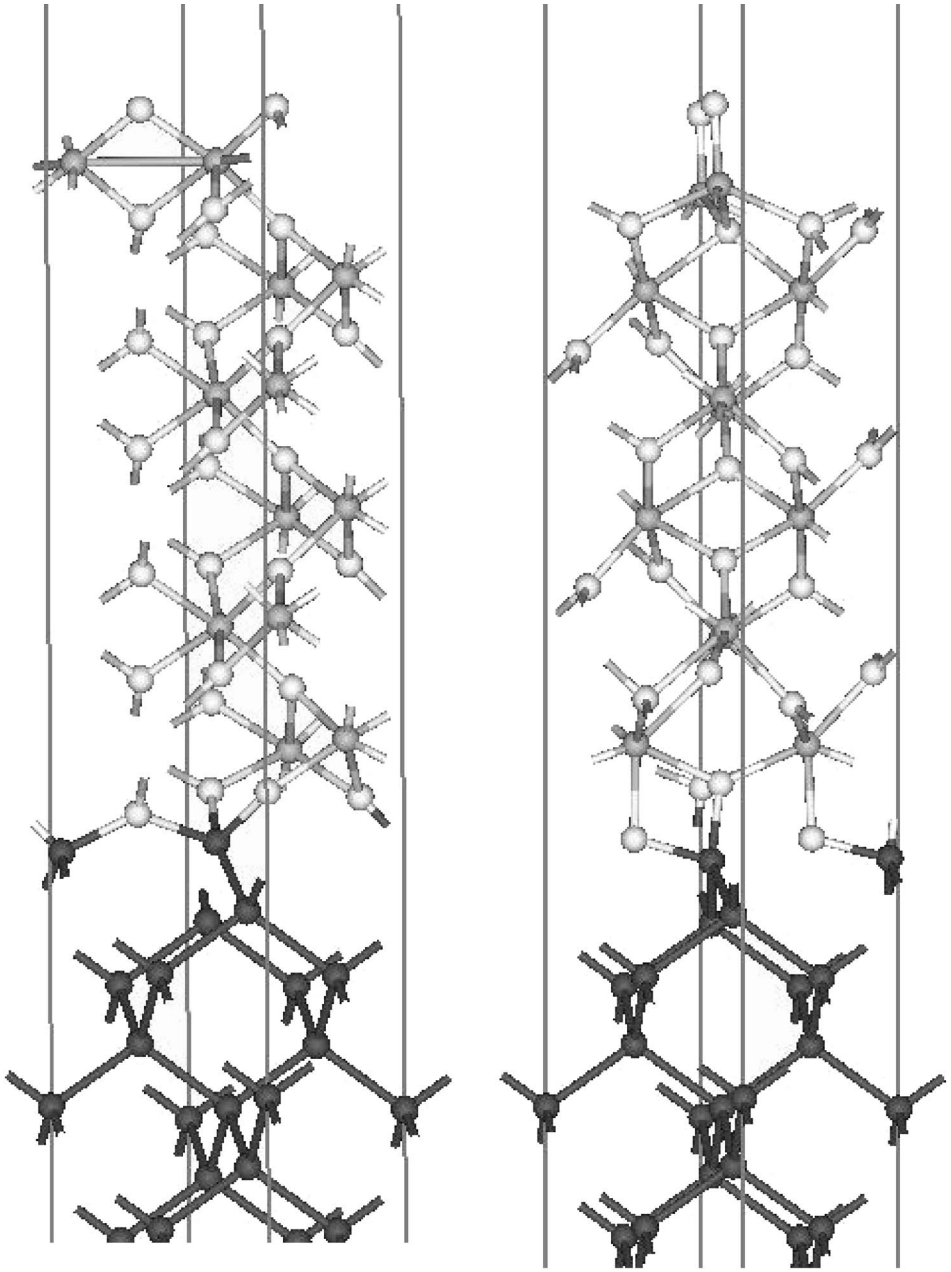

FIG. 3. Ball-and-stick models of (a) the O4 interface and (b) the O3 interface of epi-mono HfO_2 on top of Si(100) (Hf=gray, O=white, and Si=dark).

Therefore, the VBO for the O3 interface could not be reliably determined. For the O4 interface, however, only mild electric fields were present. In Fig. 4, the local potential averaged over the *x-y* plane is plotted as a function of (001) distance. The bottom ≈2.0 nm is the silicon slab, then the next ≈1.0 nm is the epi-mono HfO_2 slab, and the rest is vacuum. The (001) averaged potential in the epi-mono slab is 1.27 eV above that in silicon. In Table II, we report our VBO result of 3.1 eV between Si(001) and the O4 interface of epi-mono hafnia. Within our uncertainty, this result is identical to the value reported by Fiorentini and Gulleri[8] for the same interface. The bulk plus lineup method for determining band offsets depends on the local potential reaching a bulk value on either side of the interface. By examining Fig. 4, we can see that the local potential in HfO_2 converges after the second layer from the Si-HfO_2 interface at about 2.7 nm. Therefore, our 3.1 eV VBO result will be valid for large oxide thicknesses down to oxides including over four HfO_2 layers.

IV. DISCUSSION AND CONCLUSIONS

All present VBO results are reported in Table II. Fiorentini and Gulleri[8] add a correction to their GGA VBO result using GW band-structure calculations. Such corrections are in general appropriate and they tend to increase the VBO in silicon oxide systems because the GGA band-edge errors are larger for oxides. GW calculations are not available for our epi-model oxides. Also, oxide band edges vary significantly with phase (see column 5 of Table I), so GW corrections for one phase may differ significantly from that of another phase. Nevertheless, for Si/HfO_2 interfaces, the GW cor-

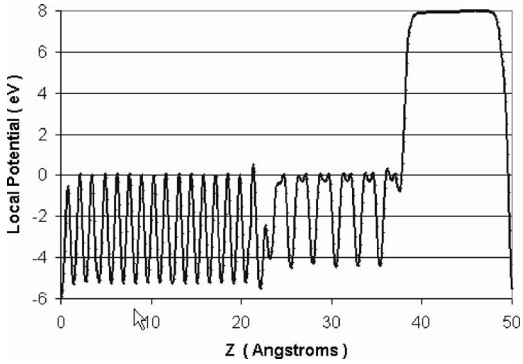

FIG. 4. The local potential averaged over the *x*-*y* plane versus the *z* coordinate for the interface model in Fig. 3(a).

rected VBO results are expected to be close to 1 eV higher than the GGA values reported in Table II.

The band diagram of Si-HfO_2 interfacial systems has been examined experimentally by several groups.[18–21] Oshima *et al.*[18] from a core-level emission spectra report a Si/HfO_2 VBO of 3.0 eV. Similarly, using x-ray photoemission, Renault *et al.*[19] report a VBO of 2.94 eV and Li *et al.*[20] report a VBO of 3.05. On the other hand Sayan *et al.*[21] using x-ray photoemission, inverse photoemission, and density-functional theory (DFT) determine a VBO of 3.25 eV when including band tail states in the analysis.

The experimental results converge to close to 3 eV for the VBO. In the present calculations, the most stable model is the epi-mono O4 interface. The VBO in this case is 3.1 eV (see Table II), in good agreement with experiments. However, as noted above, corrections to the DFT band edges will push this 3.1 eV result up by as much as 1.0 eV. Solace is found in the work of Sayan *et al.*[21] who find a larger offset (VBO=3.6 eV) if a crystalline model oxide was used in the analysis. Clearly, our models are crystalline. In actual device structures, HfO_2 is amorphous or polycrystalline.[22] Oxide bond strain disorder, grain boundaries, and point defects will all contribute to band tails. Such band tails would cause the valence band edge of HfO_2 to spread, lowering the VBO considerably. From our crystalline model oxides, it is difficult to fully access the nature of band tails in the real disordered oxides. Recent analysis of experimental electron-energy-loss spectra[21] suggests that oxygen vacancies may be important to band tail formation. Our results in Table I indicate the band gap of HfO_2 is extremely sensitive to strain, suggesting that bulk strain can also contribute to band tails. Without a good understanding of the nature of band tail states in HfO_2, it is difficult to quantify their effect on the calculated VBO. However, based on the present calculations, the GGA VBO without the GW or the band tail correction agrees with experiments. Therefore, it appears that the band tail error is a large effect on the order of the GW correction ($\sim$1 eV). Interestingly, the situation here for HfO_2 contrasts dramatically with SiO_2 whose tetragonal bonding topology is very robust. In Si/SiO_2 junctions, crystalline and amorphous oxides produce very similar VBOs.[17]

The main conclusions of this work are as follows. From the results in Tables I and II, structure, coordination, and electronic band edges are found to be very sensitive to local strain. From the results comparing O3 and O4 interfaces in Table II, the interface dipole is found to be important and sensitive to the location of the interfacial oxygen atoms. Thermodynamics favors epitaxial growth of the O4 interface. Finally, by comparing the present results to experimental studies, we find oxide band tail states to be an important determining factor in band offset formation.

ACKNOWLEDGMENTS

This work was made possible by funding from Petroleum Research Fund through a Summer Research Fellowship. The calculations were performed on the IBM SP690 machines at the National Center for Supercomputing Applications in Urbana, IL through an Alliance grant (project ngf).

*Electronic address: brt10@psu.edu

[1] Semiconductor Industry Association, International Technology Roadmap for Semiconductors, 2005 edition, Austin, TX, International Sematech, 2005. (This is available for viewing and printing from the Internet, with the following URL: http://public.ITRS.net.)

[2] T. Iwamoto *et al.*, International Electron Device Meeting Technical Digest, Washington, DC, December 2003 (unpublished), pp. 639–642.

[3] Since submission of this paper, Intel has announced their 45 nm integrated MOSFET has a hafnia-based gate dielectric.

[4] X. Zhao and D. Vanderbilt, Phys. Rev. B **65**, 233106 (2002).

[5] G. M. Rignanese, X. Gonze, G. Jun, K. Cho, and A. Pasquarello, Phys. Rev. B **69**, 184301 (2004).

[6] J. E. Jaffe, R. A. Bachorz, and M. Gutowski, Phys. Rev. B **72**, 144107 (2005).

[7] R. Ramprasad and N. Shi, Phys. Rev. B **72**, 052107 (2005).

[8] V. Fiorentini and G. Gulleri, Phys. Rev. Lett. **89**, 266101 (2002).

[9] P. W. Peacock, K. Xiong, K. Y. Tse, and J. Robertson, Phys. Rev. B **73**, 075328 (2006).

[10] R. Puthenkovilakam, E. A. Carter, and J. P. Chang, Phys. Rev. B **69**, 155329 (2004).

[11] Y. F. Dong, Y. P. Feng, S. J. Wang, and A. C. H. Huan, Phys. Rev. B **72**, 045327 (2005).

[12] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993); M. Ueno, M. Yoshida, A. Onodera, O. Shimomura, and K. Takemura, *ibid.* **49**, 14251 (1994).

[13] G. Kresse and J. Hafner, J. Phys.: Condens. Matter **6**, 8245 (1994).

[14] J. P. Perdew and Y. Wang, Phys. Rev. B **46**, 12947 (1992).

[15] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[16] C. G. Van de Walle and R. M. Martin, Phys. Rev. B **35**, 8154 (1987).

[17] B. R. Tuttle, Phys. Rev. B **70**, 125322 (2004).

[18] M. Oshima, S. Toyoda, T. Okumura, J. Okabayashi, H. Kumigashira, K. Ono, M. Niwa, K. Usuda, and N. Hirashita, Appl. Phys. Lett. **83**, 2172 (2003).

[19] O. Renault, N. T. Barrett, D. Samour, and S. Quiais-Marthon, Surf. Sci. **566-568**, 526 (2004).
[20] Q. Li, S. J. Wang, K. B. Li, A. C. Huan, J. W. Chai, J. S. Pan, and C. K. Ong, Appl. Phys. Lett. **85**, 6125 (2004).
[21] S. Sayan, T. Emge, E. Garfunkel, X. Zhao, L. Wielunski, R. Bartynski, D. Vanderbilt, J. Suehle, S. Suzer, and M. Banaszak-Holl, J. Appl. Phys. **96**, 7485 (2004).
[22] X. Baik *et al.*, Appl. Phys. Lett. **85**, 672 (2004).