Journal of Physics Communications

PAPER • **OPEN ACCESS**

GGA-1/2 self-energy correction for accurate band structure calculations: the case of resistive switching oxides

To cite this article: Jun-Hui Yuan *et al* 2018 *J. Phys. Commun.* **2** 105005

IOP Publishing J. Phys. Commun. 2 (2018) 105005 https://doi.org/10.1088/2399-6528/aade7e

Journal of Physics Communications

OPEN ACCESS

RECEIVED
5 June 2018

REVISED
15 August 2018

ACCEPTED FOR PUBLICATION
3 September 2018

PUBLISHED
4 October 2018

PAPER

GGA-1/2 self-energy correction for accurate band structure calculations: the case of resistive switching oxides

Jun-Hui Yuan[1], **Qi Chen**[1], **Leonardo R C Fonseca**[2], **Ming Xu**[1], **Kan-Hao Xue**[1] **and Xiang-Shui Miao**[1]

[1] Wuhan National Research Center for Optoelectronics, School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China

[2] Departamento de Física, ICEx, Universidade Federal de Minas Gerais, 30123-970 Belo Horizonte, MG, Brazil

E-mail: xkh@hust.edu.cn

Keywords: GGA-1/2, density functional theory, band gap, self-energy correction, resistive switching oxides

Abstract

First-principles calculation has become an indispensable methodology in revealing the working principles of nanoscale electronic devices, but ultra-large supercells are usually required in modeling the devices with critical metal/dielectric interfaces. Traditional density functional theory within the generalized gradient approximation (GGA) suffers from the inaccurate band gap problem when metal oxides are present, but they serve as the core component in resistive random access memory (RRAM), which is a promising path for novel high speed non-volatile memories. To obtain improved oxide band gaps, we applied the efficient GGA-1/2 method for self-energy correction, whose computational load is at the same level as standard GGA. In particular, we have investigated the influence of exchange-correlation functional flavors on the GGA-1/2 band structures, taking four important binary oxide RRAM materials (α-Al_2O_3, r-TiO_2, m-ZrO_2 and m-HfO_2) as benchmark examples. Five GGA functionals (PBE, PBEsol, PW91, revPBE and AM05) were considered and their band structures were compared in detail. We have found that the performance of GGA-1/2 is comparable to state-of-the-art *GW* and generally superior to the HSE06 hybrid functional. Among the five GGA functionals, PBEsol yields the best results in general. In addition, the applicability of a single self-energy potential for various GGA-1/2 flavors is discussed. Our work provides a guide to the GGA flavor selection, when applying the GGA-1/2 method to metal oxides.

1. Introduction

With the rising of big data [1], cloud computing [2], internet of things [3] and other technologies in recent years, an exploding amount of information requires storage and analysis, for which fast, low power consumption, high density and non-volatile memories are suitable. However, the NAND FLASH memory that dominates the present non-volatile memory market suffers from intrinsically slow erasing operation, opening an urgent demand for next-generation non-volatile memories [4, 5]. Resistive random access memory (RRAM), which achieves data storage through reversible resistance changes in certain dielectrics, is one of the leading candidates for this purpose. Reproducible resistive switching phenomena have been widely observed in oxides [6–9], chalcogenides [10] and even organic materials [11]. Due to their simple device structure (metal/insulator/metal), high compatibility to standard CMOS flow line and the prospect of 3D integration, certain binary oxides such as Al_2O_3 [12], HfO_2 [13, 14], TiO2 [15, 16], and ZrO_2 [17] have attracted more attention. In recent few years, HfO2 has been regarded in some works [18] as the most promising candidate material for commercial application.

Although oxide-based RRAM has experienced a rapid development for more than a decade, it has not reached large-scale industrial applications yet. There are still some fundamental issues to be resolved in terms of the resistive switching mechanism, reliability, integration schemes, and so forth [15]. For instance, the mesoscopic physical mechanisms of switching and the corresponding models, without which a proper device optimization cannot be achieved, still require more in-depth study. Most binary metal oxides are believed to manifest a transition from a high resistance state (HRS) to a low resistance state (LRS) upon the formation of conductive filaments (CFs) connecting the two electrodes. Yet, at present there are still controversies on the exact composition and geometry of the CFs [19–22]. Since it is difficult to directly observe and determine experimentally the exact composition of the CFs within the dielectric media, theoretical approaches can be particularly helpful.

Among atomistic simulation techniques, *ab initio* calculations based on density functional theory (DFT) [23] are highly successful in the prediction of total energies, bond lengths, and vibration frequencies. However, DFT within local density approximation (LDA) or generalized gradient approximation (GGA) severely underestimates the band gaps of semiconductors and insulators, which limits its capability of predicting the band alignment in metal/semiconductor contacts, a significant task in understanding charge transmission through CFs [24]. Various theoretical approaches have been proposed to overcome this shortcoming, thus providing more reliable predictions of physical properties that depend on excited states. Hedin's *GW* approximation is considered to be the state-of-the-art, which is still the most accurate method for electronic structure calculation. It calculates the quasiparticle energies in terms of the perturbation theory [25], and is a beyond-DFT technique. Hybrid functionals [26], which combine standard DFT and Hartree–Fock exchange functionals through a mixing parameter, are also well-known and robust alternatives, though they sometimes require an adjustment of the mixing parameter. Both hybrid functional and *GW* methods demand much greater computational efforts than standard DFT-LDA or DFT-GGA, setting severe limitations to the size of the simulated systems. Indeed, to simulate realistic RRAM devices, large supercells with hundreds of atoms are typically required. For instance, to investigate the atomic structure of a CF through the dielectric, a large cross-section area is indispensable to avoid undesired interactions between a CF and its periodic images [13]. Therefore, realistic *ab initio* RRAM device calculations require both a high computational speed and relatively accurate band gaps. Traditionally, one resolves the band gap problem in DFT-LDA/DFT-GGA by employing LDA + U and GGA + U [27–30], or resorts to hybrid functionals [31]. While the adoption of hybrid functional increases the computational load considerably, LDA + U is almost as fast as conventional LDA. However, LDA + U is not suitable for many of these RRAM oxides. For example, Al_2O_3, TiO_2, ZrO_2 and HfO_2 are all ordinary oxides without strong electron correlation, in which case LDA + U does not work well unless an additional large U_p correction to oxygen is considered [27], reflecting the fact that the band gap inaccuracy lies primarily with the oxygen anion rather than the metal cation.

The LDA-1/2 method corrects the spurious electron self-interaction term in LDA using the Slater half-occupation technique [32]. As the conduction band electron usually occupies Bloch-like states with nearly-null self-energy, corrections are only carried out to the holes that are typically localized on the anions. Instead of calculating explicitly the self-energy of the hole, such correction can be achieved by simply introducing a so-called self-energy potential in real space, which acts as an external potential to the region where the hole resides. The exact form of self-energy potential for any atom can be derived *ab initially* from atomic calculations using DFT, with the desired exchange-correlation (XC) flavor. Such self-energy potential for the anion is then trimmed to a finite range to avoid overlapping between neighboring anions. Going one step further from LDA-1/2, a natural extension is GGA-1/2, which was previously shown to bring about improvement over LDA-1/2 in some cases [33]. While LDA tends to underestimate the lattice constants, in the framework of GGA there are certain XC functionals like PBEsol [34], Wu-Cohen [35] and AM05 [36] that can predict rather accurate lattice constants. Hence, GGA-1/2, which corrects the electronic band structure calculated with GGA, ought to be a better computational method for RRAM simulation than LDA-1/2. However, to our best knowledge there is no theoretical study on the influence of GGA flavors towards the accuracy of GGA-1/2 band structures. Hence, from both theoretical and application aspects, it is of great significance to carry out a systematic study on the GGA-1/2 band structures built upon various GGA flavors.

In this work, we adopt four relevant binary oxide RRAM materials as benchmarks. Five GGA XCs are considered and their band structures are compared in detail. Our results offer guidance in properly choosing GGA flavors to achieve efficient and accurate GGA-1/2 band structure calculations for the four binary oxide semiconductors, i.e. corundum Al_2O_3, rutile TiO_2, monoclinic ZrO_2 and monoclinic HfO_2. Our results can be readily extended to more oxides.

2. Computational

2.1. Crystal structures

We studied the room temperature phases of Al_2O_3, TiO_2, ZrO_2 and HfO_2, namely corundum (α-Al_2O_3), rutile-TiO_2 (r-TiO_2), monoclinic ZrO_2 (m-ZrO_2) and monoclinic HfO_2 (m-HfO_2), shown in figure 1. Corundum α-Al_2O_3 crystallizes with trigonal symmetry in the space group $R\bar{3}c$, where oxygen atoms are arranged by a slightly distorted hexagonal close packing, in which two-thirds of the gaps between the octahedra

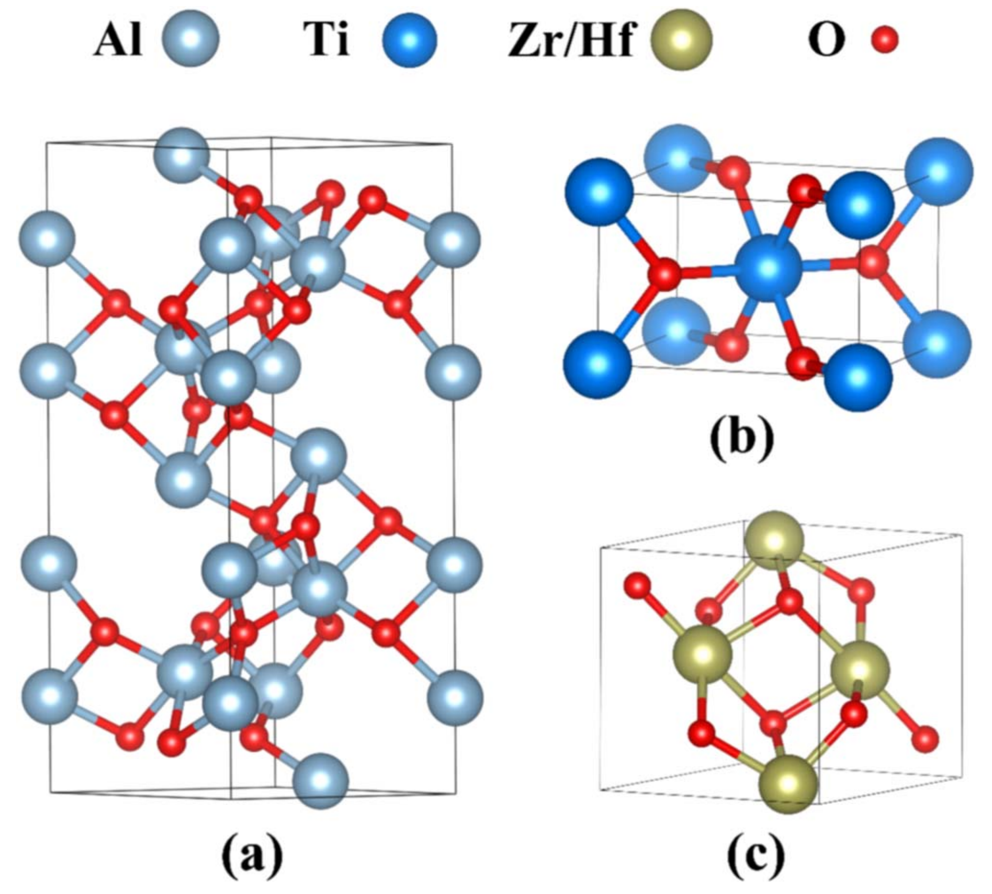

Figure 1. Crystal structures of (a) α-Al_2O_3, (b) r-TiO_2, (c) isomorphic m-ZrO_2 and m-HfO_2.

are occupied by aluminum (figure 1(a)) [37]. Rutile TiO_2 presents a tetragonal unit cell with the titanium atoms surrounded by an octahedron of six oxygen atoms. The oxygen coordination number is 3, resulting in a trigonal planar geometry, as illustrated in figure 1(b) [38]. The ground state monoclinic phases of ZrO_2 and HfO_2 are isomorphic, both with the baddeleyite structure ($P2_1/c$) as shown in figure 1(c) [39]. Unlike TiO_2, which features six-coordinated Ti in all phases, m-ZrO_2 and m-HfO_2 consist of seven-coordinated zirconium or hafnium centers. This discrepancy is attributed to the larger size of Zr/Hf atoms relative to the Ti atom.

2.2. Computational methods

DFT calculations were carried out using the plane-wave-based Vienna *Ab initio* Simulation Package (VASP) [40, 41], with a fixed 500 eV plane-wave kinetic energy cutoff. Five distinct flavors for the GGA XC functional were considered: Perdew–Burke–Ernzerhof (PBE) [42], PBE developed for solids (PBEsol) [34], Perdew-Wang 1991 (PW91) [43], the revised PBE (revPBE) [44] and the Armiento–Mattsson 2005 (AM05) [36]. In addition, we employed the HSE06 screened hybrid functional [26] (using the PBE-optimized geometries) to calculate the band structures of all the four crystals for comparison. The HSE06 functional used in this work assumes its most common form, constructed by mixing 25% of the Fock exchange with 75% of the PBE exchange, and 100% of the PBE correlation [26]. Core electrons were approximated by projector augmented-wave pseudopotentials [45, 46]. Unless specified otherwise, the electrons considered as valence were: 3s and 3p for Al; 3s, 3p, 3d and 4s for Ti; 4s, 4p, 4d and 5s for Zr; 5p, 5d and 6s for Hf; 2s and 2p for O. In all self-consistent runs, the total energy convergence criterion was set to below 10^{-6} eV, while structural optimization was reached for Hellmann-Feynman forces less than 0.01 eV/Å in any direction. We employed the following Monkhorst-Pack k-point grids for Brillouin zone sampling during geometry optimization: $5 \times 5 \times 2$ Γ-centered for Al_2O_3, $7 \times 7 \times 11$ for TiO_2, $7 \times 7 \times 7$ for ZrO_2, and $7 \times 6 \times 6$ for HfO_2. For total energy calculations the k-point grids were enlarged to $10 \times 10 \times 4$ for Al_2O_3, $11 \times 11 \times 17$ for TiO_2, $10 \times 10 \times 10$ for ZrO_2, and $11 \times 9 \times 9$ for HfO_2 [47].

2.3. GGA-1/2 for improved band structures

GGA-1/2 generalizes to solids the original idea of Slater's half occupation scheme, which was initially designed for ionization energy calculation in isolated atoms. Its formalism starts from Janak's formula, which offers an interpretation of the DFT single particle eigenvalues e_α as the variation of the total energy E with respect to its occupation f_α

$$\frac{\partial E}{\partial f_\alpha} = e_\alpha(f_\alpha) \tag{1}$$

Assuming a linear relation between the eigenvalue and its occupation [48], it is straightforward to show that

$$E(0) - E(-1) = e_\alpha(0) - S_\alpha \tag{2}$$

Equation (2) means that the energy difference between the ground state energy $E(0)$ and the total energy of the ion $E(-1)$ with one electron removed from state α, is given by the energy eigenvalue of state α and a so-called "self-energy term" S_α. The analytical expression for S_α [32],

$$S_\alpha = \iint d^3r d^3r' \frac{n_\alpha(r) n_\alpha(r')}{|r - r'|} + higher\ order\ terms \tag{3}$$

indeed has the form of a self-energy, which renders its name. If one regards S_α as the quantum mechanical average of a "self-energy potential" V_s,

$$S_\alpha = \int d^3r n_\alpha(\mathbf{r}) V_s(\mathbf{r}) \tag{4}$$

it follows that V_s can be obtained as the difference from the DFT-calculated electrostatic potentials between the neutral atom and the half-ionized ions:

$$V_s \approx -V(-1/2, r) + V(0, r) \tag{5}$$

As mentioned in section 1, the self-energy potential must be trimmed by some function $\Theta(r)$ [32]. To obtain the rectified band gaps for solid state materials, V_s is added to the pseudopotentials of the anions and a second self-consistent run using the modified pseudopotentials is carried out upon the same relaxed geometry. Since the self-energy potentials serve as external potentials in the solid, the ground state energies given by GGA-1/2 self-consistent calculations are not physically meaningful themselves. Nevertheless, band diagrams and density of states are properly recovered by GGA-1/2 [32]. Hence, in our work all total energies were taken from GGA calculations, while the band diagrams were taken either from GGA (uncorrected) or from GGA-1/2 (corrected) calculations.

Since all the materials under investigation are binary metal oxides, the self-energy correction ought to be done only for oxygen anions. The optimal cutoff radius in $\Theta(r)$ to trim the corresponding V_s (with (1/2)e subtracted from the O 2p orbital [32]) was obtained variationally upon maximizing the band gap [32]. Electronic structures were obtained with this optimal cutoff radius only, and no empirical parameters were involved.

The self-energy potentials with all these GGA XC flavors (PBE, PBEsol, PW91, revPBE, AM05) were derived from atomic calculations using a modified ATOM code (supplied with the Siesta simulation package [49]). These self-energy potentials do not depend on the *ab initio* simulation code, and can be regarded as only specific to the XC flavor. We have also created a web-based self-energy correction program where pseudopotentials modified through the inclusion of GGA-1/2 self-energy potentials for all these XC flavors can be generated online, at http://www.eedevice.com/dft-half.

3. Results and discussions

3.1. Geometric structures and partial density of states

The optimized lattice parameters for these four crystals are summarized in table 1, which also shows the corresponding experimental values [50–52]. Due to their intrinsic similarity, PBE and PW91 yield very similar lattice parameters, while the lattice parameters obtained with PBEsol and AM05 also share a similar trend. The lattice parameters obtained with revPBE are always the largest among all considered XC functionals. Compared with experimental values, PBEsol and AM05 are closest to data, while PBE and PW91 usually predict slightly larger lattice constants. On the other hand, revPBE yields the worst agreement of all functionals tested.

The partial density of states for α-Al_2O_3, r-TiO_2, m-ZrO_2 and m-HfO_2 obtained with GGA-PBE are shown in figure 2. In any material the valence band maximum (VBM) mainly consists of O-2p states, which confirms that the hole is located around oxygen anions and the self-energy correction should indeed be done to oxygen only.

3.2. Band structure results based on optimized lattice parameters

The band gaps calculated with PBE, PBEsol, PW91, revPBE, AM05, HSE06 and GGA-1/2 (including the five different XC flavors) for the four compounds are listed in table 2, where fully optimized lattice parameters were utilized for each XC. As expected, DFT-GGA severely underestimates the energy gap. All band gaps obtained with conventional GGA are far below the experimental values by 1.16 eV $\sim$ 2.67 eV. Among the four materials,

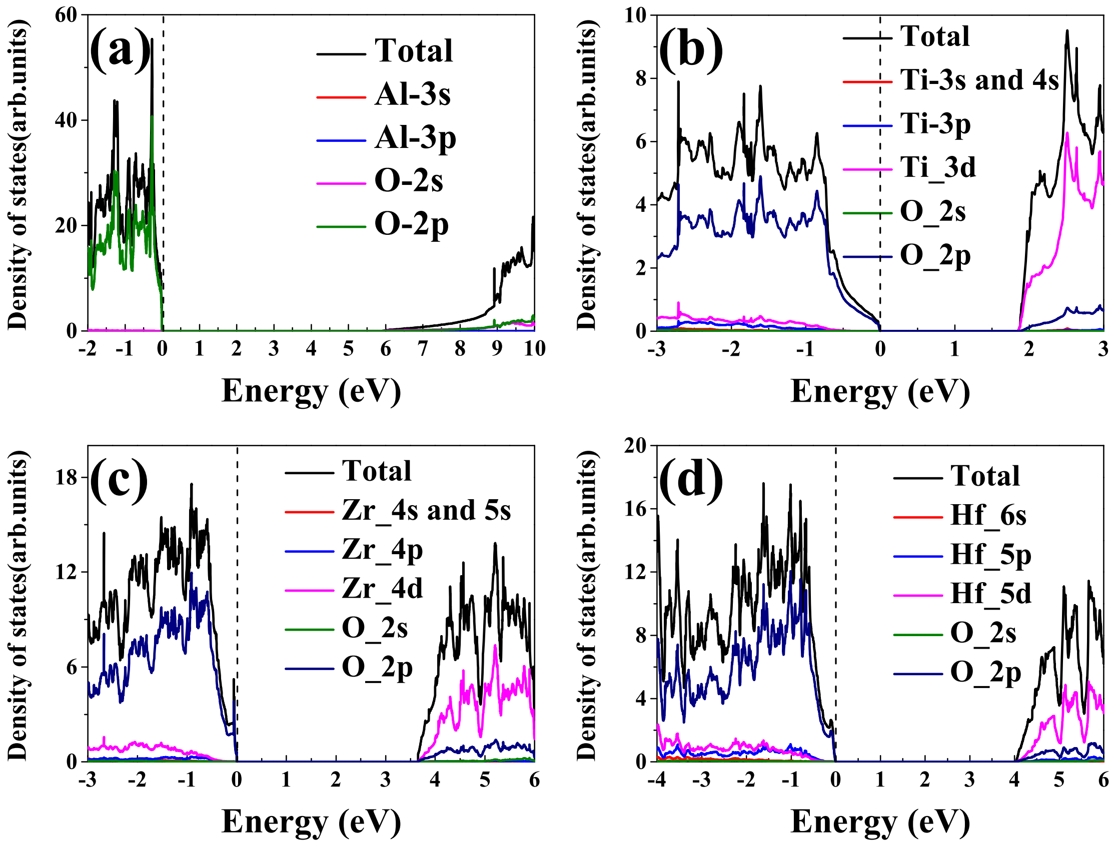

Figure 2. Partial density of states for (a) α-Al_2O_3, (b) r-TiO_2, (c) m-ZrO_2 and (d) m-HfO_2, obtained with conventional GGA-PBE.

Table 1. Optimized structural parameters for α-Al_2O_3, r-TiO_2, m-ZrO_2 and m-HfO_2, calculated using various XC functionals. Experimental values are shown for comparison.

	PBE	PBEsol	PW91	revPBE	AM05	Experimental
			α-Al_2O_3			
a(Å)	4.807	4.774	4.804	4.837	4.775	4.759 [50]
c(Å)	13.122	13.013	13.111	13.211	13.020	12.991
			r-TiO_2			
a(Å)	4.644	4.593	4.644	4.687	4.600	4.601 [51]
c(Å)	2.968	2.942	2.968	2.984	2.941	2.968
			m-ZrO_2			
a(Å)	5.187	5.132	5.213	5.227	5.134	5.146 [52]
b(Å)	5.247	5.212	5.279	5.273	5.214	5.208
c(Å)	5.376	5.301	5.395	5.427	5.305	5.315
β angle (°)	99.602	99.596	99.603	99.678	99.658	99.249
			m-HfO_2			
a(Å)	5.138	5.079	5.137	5.180	5.083	5.118 [52]
b(Å)	5.191	5.156	5.190	5.218	5.152	5.186
c(Å)	5.321	5.246	5.320	5.369	5.253	5.284
β angle (°)	99.697	99.664	99.680	99.740	99.759	99.352

α-Al_2O_3 is a wide band gap insulator with the VBM and conduction band minimum (CBM) both located at Γ, which is consistent with previous results [53, 54]. The largest obtained band gap for α-Al_2O_3 based on DFT-GGA is 6.13 eV, using the AM05 XC. However, that is still 2.67 eV smaller than the typical experimental value (8.8 eV) [53], and even smaller than the previous *GW* results (9.36 eV and 9.78 eV) [55, 56]. After self-energy correction, the GGA-1/2 band gaps obtained with the five XC functionals are all above 8 eV (table 2), which is comparable to the HSE06 values 8.09 eV (calculated previously for a PBEsol-optimized structure using norm-conserving pseudopotentials [54]) and 7.70 eV (calculated in this work for a PBE-optimized structure).

Table 2. Calculated GGA and GGA-1/2 band gaps (unit: eV) for α-Al_2O_3, r-TiO_2, m-ZrO_2 and m-HfO_2 with various exchange-correlation functionals, based upon optimized lattice parameters. HSE06, *GW* and experimental values are also listed for comparison.

	PBE	PBEsol	PW91	revPBE	AM05	HSE06	*GW*	Expt.
					α-Al_2O_3			
GGA	5.86	6.04	5.93	5.72	6.13	7.70,	9.36 [55],	8.8 [53]
GGA-1/2	8.10	8.25	8.16	8.04	8.38	8.088[a] [54]	9.78 [56]	
					r-TiO_2			
GGA	1.82	1.84	1.78	1.83	1.84	3.32,	3.46(G_0W_0) [57]	3.0–3.4
GGA-1/2	3.04	3.10	2.97	3.02	3.09	3.39 [57]		[58–61]
					m-ZrO_2			
GGA	3.63	3.58	3.48	3.70	3.61	5.20,	4.99(G_0W_0),	5.0–5.8 [62],
GGA-1/2	5.51	5.45	5.33	5.59	5.50	5.14 [63]	5.34(GW_0) [62], 5.42(GW_0) [64]	5.4 [65]
					m-HfO_2			
GGA	4.03	3.96	4.01	4.10	3.99	5.56,	6.0 [66],	5.9 [67, 68]
GGA-1/2	6.06	5.96	6.04	6.17	6.02	5.83[a] [69]	5.45(G_0W_0), 5.78(GW_0) [62], 5.92(G_0W_0) [69]	

[a] With norm-conserving pseudopotentials

Although the values are still smaller than the experimental values by about 0.4 eV ~ 0.8 eV (below we show that α-Al_2O_3 is the most difficult case among the materials under investigation), they do show great improvement over conventional DFT-GGA. We have noticed that the band gaps of α-Al_2O_3 obtained with HSE06 are about 0.7 eV ~ 1.1 eV smaller than experimental, while the *GW* band gaps are 0.6 eV ~ 1.0 eV larger than experimental. Our GGA-1/2 band gaps fall between HSE06 and *GW* results.

For *r*-TiO_2, all the calculated band structures exhibit a direct band gap with the VBM and CBM both located at Γ, consistent with previous experimental and theoretical results [57–61]. Compared with experimental band gap values [58–61], the GGA band gaps differ by about 1.16 eV ~ 1.56 eV (table 2). Yet, the GGA-1/2 band gaps are quite satisfactory, agreeing well with experimental values by showing merely −0.1 eV to 0.4 eV mismatches. In addition, our GGA-1/2 results are also very close to the HSE06 (3.32 eV and 3.39 eV [58]) and *GW* (3.46 eV) [58] results.

Regardless of the computational method, both *m*-ZrO_2 and *m*-HfO_2 are predicted to be indirect band gap semiconductors with the VBM and CBM located at Γ and B, respectively. In general, the calculated GGA band gaps are lower than typical experimental values by about 1.30 eV ~ 2.32 eV (for *m*-ZrO_2) and 1.80 eV ~ 1.94 eV (for *m*-HfO_2). After self-energy correction, the GGA-1/2 band gaps differ from the experimental value by merely −0.59 eV ~ 0.47 eV for *m*-ZrO_2 and −0.27 eV ~ −0.06 eV for *m*-HfO_2, respectively. Regarding higher order calculations, the HSE06 band gap for *m*-ZrO_2 (5.20 eV in this work and 5.14 eV from [63]) is lower than our GGA-1/2 results as well as the available experimental values. Moreover, our GGA-1/2 results are comparable to that from GW_0 (5.34 eV [62] and 5.42 eV [64]) and larger than that of G_0W_0 (4.99 eV) calculations [62]. Comparing the five XC flavors, PBEsol yields the best band gap that differs from experimental by less than 0.05 eV (adopting the experimental value of 5.4 eV [65]), followed by PW91 (0.07 eV), AM05 (0.10 eV), PBE (0.11 eV) and revPBE (0.19 eV).

For *m*-HfO_2, the calculated self-energy corrected band gaps are nearly in line with the experimental results (5.9 eV band gap [67, 68]), where the largest difference 0.27 eV comes from the revPBE XC. PBEsol yields the best agreement (differing by 0.06 eV), as it does for *m*-ZrO_2, followed by AM05 (0.12 eV), PW91 (0.14 eV), PBE (0.16 eV) and revPBE (0.27 eV). In addition, the values obtained with HSE06 (5.56 eV and 5.83 eV [69]) are slightly smaller than experimental. Notice that the *m*-HfO_2 band gaps obtained with *GW* range from 5.45 eV to 6.0 eV [62, 66, 69], similar to the GGA-1/2 range of values obtained with any of the five GGA-XC functionals.

Figure 3 shows the best (among all XC functionals) GGA-1/2 band structures for each material, along with conventional DFT-GGA and HSE06 results. The GGA-1/2 band structures are in line with DFT-GGA and HSE06 in terms of the morphology and the intrinsic excitation characteristics (such as the locations of VBM and CBM). This confirms that for these oxides, the good accuracy of GGA-1/2 is not limited to the fundamental gap, but also for the overall band structures, where GGA-1/2 is comparable with computationally intensive methods such as hybrid functionals.

3.3. Band structure results based on experimental lattice parameters

Since each GGA XC leads to a specific set of optimum lattice parameters, the band diagrams calculated in section 3.2 were not based on a unique unit cell for each material. To better evaluate the difference among various XCs, we further calculated the band gaps of α-Al_2O_3, *r*-TiO_2, *m*-ZrO_2 and *m*-HfO_2 with both GGA and GGA-1/2 using their experimental lattice parameters. Yet, to avoid internal forces, the atomic coordinates within the unit cell were fully relaxed. The calculated band gaps for all five XC flavors are listed in table 3, and figure 4(a) compares those band gaps obtained with optimized lattice parameters with fixed experimental lattice parameters. Among the five XCs, the obtained band gap of *r*-TiO_2 with experimental lattice parameters is slightly smaller than that of optimized structure for PBEsol, PW91 and AM05, regardless of using GGA or GGA-1/2. On the other hand, the results of PBE (both GGA and GGA-1/2) and revPBE (only GGA-1/2) are slightly larger than that of optimized structure. In addition, for α-Al_2O_3 the band gaps are much improved by fixing to experimental lattice constants, especially for the revPBE XC whose GGA-1/2 band gap now only deviates from the experimental value by 0.15 eV. It is therefore clear that the original poor GGA-1/2 band gap using equilibrium lattice constants is strongly related to the over-estimation of lattice constants. For instance, the equilibrium *c*-axis lattice constant was predicted to be 13.211 Å using the revPBE XC, which is 1.7% larger than the experimental value. Such improvement also implies that using correct lattice constants is more important than minimizing the stress, in order to obtain band gaps close to measured values. Finally, we notice that the lattice constant effect on *m*-ZrO_2 and *m*-HfO_2 is relatively tiny, because the band gaps derived from either optimized lattice or experimental lattice differ by no more than 0.1 eV.

It is well-known that AM05 and PBEsol yield the most accurate lattice parameters among common GGA XCs. Therefore, it is natural to infer that for these two XCs the GGA-1/2 band gaps do not vary much when switching to experimental lattice parameters. However, the quantitative relation is still unclear yet. Hence, we have also listed explicitly in figure 4(b) the band gap shift [$\Delta E_g = E_g$ (experimental lattice) - E_g (optimized

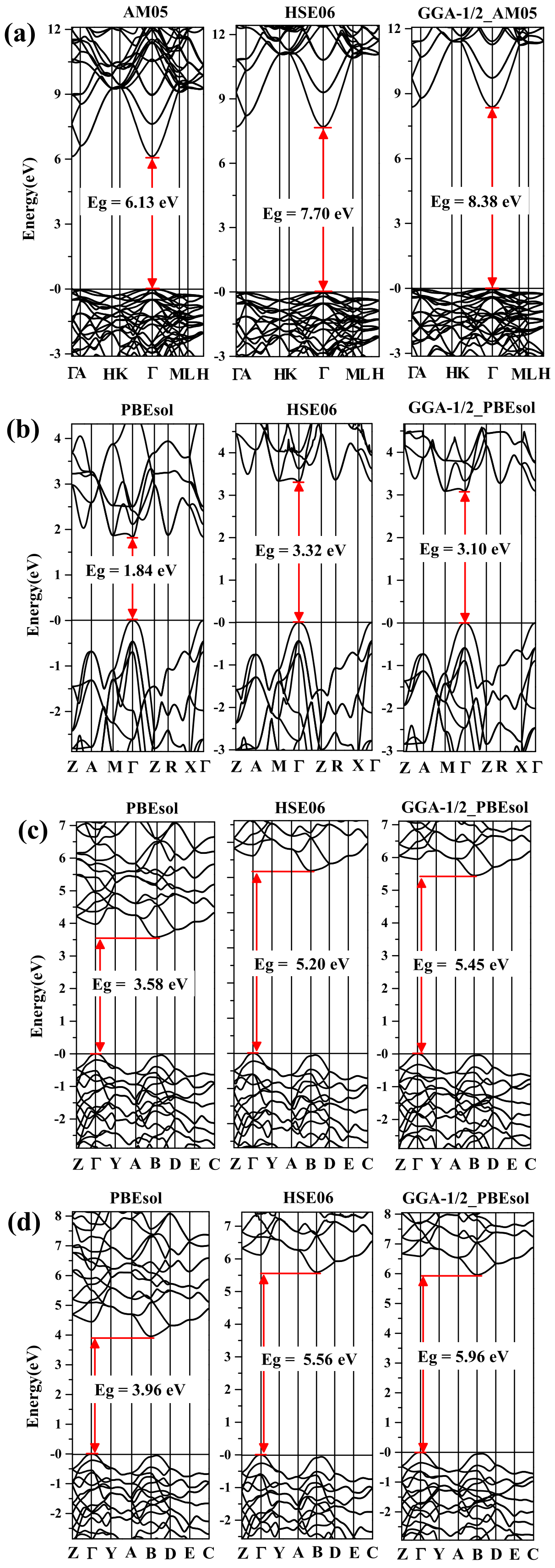

Figure 3. (a) Band diagrams of α-Al_2O_3 calculated with GGA-AM05, HSE06 and GGA-1/2 (AM05); (b) band diagrams of *r*-TiO_2 calculated with GGA-PBEsol, HSE06 and GGA-1/2 (PBEsol); (c) band diagrams of *m*-ZrO_2 calculated with GGA-PBEsol, HSE06 and GGA-1/2 (PBEsol); (d) band diagrams of *m*-HfO_2 calculated with GGA-PBEsol, HSE06 and GGA-1/2 (PBEsol).

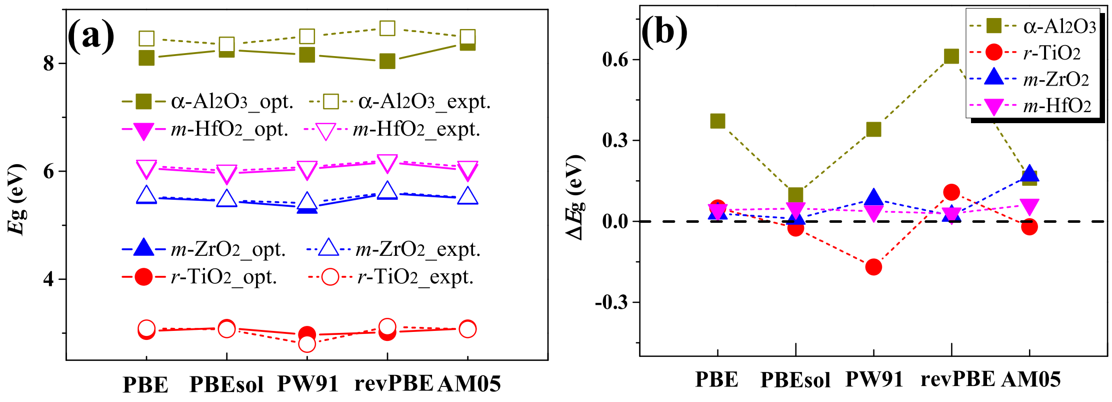

Figure 4. (a) Energy band gaps and (b) differences of the band gaps for α-Al_2O_3, r-TiO_2, m-ZrO_2 and m-HfO_2 calculated by GGA-1/2, based on either optimized or experimental lattice parameters.

Table 3. Calculated GGA and GGA-1/2 band gaps (unit: eV) for α-Al_2O_3, r-TiO_2, m-ZrO_2 and m-HfO_2 with various exchange-correlation functionals, based upon experimental lattice parameters. HSE06, *GW* and experimental values are also listed for comparison.

	PBE	PBEsol	PW91	revPBE	AM05	HSE06	*GW*	Expt.
					α-Al_2O_3			
GGA	6.22	6.13	6.26	6.31	6.23	8.088[a] [54]	9.36 [55], 9.78 [56]	8.8 [53]
GGA-1/2	8.46	8.35	8.50	8.65	8.49			
					r-TiO_2			
GGA	1.83	1.81	1.66	1.78	1.82	3.39 [57]	3.46(G_0W_0) [57]	3.0–3.4 [58–61]
GGA-1/2	3.09	3.07	2.80	3.12	3.07			
					m-ZrO_2			
GGA	3.65	3.59	3.57	3.70	3.63	5.14 [63]	4.99(G_0W_0), 5.34(GW_0) [62], 5.42(GW_0)[64]	5.0–5.8 [62], 5.4 [65]
GGA-1/2	5.53	5.46	5.41	5.61	5.51			
					m-HfO_2			
GGA	4.08	4.01	4.06	4.13	4.06	5.83[a] [69]	6.0 [66], 5.45(G_0W_0), 5.78(GW_0) [62], 5.92(G_0W_0) [64]	5.9 [67, 68]
GGA-1/2	6.10	6.01	6.08	6.20	6.08			

[a] With norm-conserving pseudopotentials.

lattice)] using GGA-1/2. Notice that while the differences for m-ZrO_2 and m-HfO_2 are slight regardless of the XC used, for α-Al_2O_3 and r-TiO_2 the differences can be a lot more pronounced. Moreover, among the five GGA XCs the ΔE_g values obtained with PBEsol have the least fluctuation in general, slightly better than AM05 (both PBEsol and AM05 are considerably better than the other three GGA XCs).

3.4. Influence of different pseudopotential forms

Pseudopotentials of different forms can have great influence on the calculated electronic structures, especially for transition metals due to the complexity of d electrons. The pseudopotential of Al is sufficiently simple because only 3 electrons need to be considered as in the valence. On the other hand, for the 4d metal Zr, the pseudopotential of the Zr_sv form (including 4s, 4p, 4d and 5s electrons in the valence) is unavoidable in VASP. There is thus little room to test the influence of pseudopotential forms for Al and Zr. Hence, we only considered different GGA-PBEsol pseudopotential forms for Ti and Hf available with the VASP code, which are Ti(Hf) (four valence electrons, 4e for short), Ti(Hf)_pv (ten valence electrons, 10e for short), and Ti(Hf)_sv (twelve valence electrons, 12e for short). As shown in table 4, the calculated band gap of r-TiO_2 is much more sensitive to the choice of pseudopotential than that of m-HfO_2. For r-TiO_2, at least the Ti_pv (10e) pseudopotential should be used in order to get accurate results. In addition, all the three pseudopotential forms considered here for Hf are suitable for various calculations. Thus, on account of both computational accuracy and efficiency, in carrying out GGA-1/2 electronic structure calculations we recommend Ti_sv (12e) or at least Ti_pv (10e) pseudopotentials for r-TiO_2, while for m-HfO_2 all three available pseudopotential forms are suitable.

Table 4. Comparison of calculated band gaps (unit: eV) for r-TiO$_2$ and m-HfO$_2$ with different versions of pseudopotentials. The GGA-1/2 method is used, within the PBEsol functional.

	Ti(Hf) (4e)	Ti(Hf)_pv (10e)	Ti(Hf)_sv (12e)
r-TiO$_2$	2.81	3.05	3.10
m-HfO$_2$	5.88	5.96	5.94

Table 5. Calculated GGA-1/2 band gaps (unit: eV) for α-Al$_2$O$_3$, r-TiO$_2$, m-ZrO$_2$ and m-HfO$_2$ with five different exchange-correlation functionals. Here the self-energy potentials were nevertheless only of PBE/PBEsol flavors.

	PBE	PBEsol	PW91	revPBE	AM05
α-Al$_2$O$_3$	8.09/8.11	8.24/8.25	8.15/8.16	8.05/8.07	8.37/8.39
r-TiO$_2$	3.04/3.04	3.09/3.09	3.01/3.02	3.02/3.03	3.08/3.09
m-ZrO$_2$	5.51/5.51	5.44/5.45	5.38/5.39	5.61/5.61	5.49/5.50
m-HfO$_2$	6.06/6.06	5.95/5.96	6.04/6.04	6.18/6.18	6.01/6.01

3.5. Universality of the GGA-1/2 self-energy potentials

In all the GGA-1/2 calculations above, we have derived the self-energy potentials by subtracting the potential of a charged ion with (1/2)e missing, from that of a neutral atom. Both neutral and ionized atomic potentials were calculated *ab initially* using DFT-GGA for each XC flavor. These self-energy potentials in fact do not differ much since the subtraction may eliminate most of the specific features belonging to each specific XC flavor. To investigate whether a unique self-energy potential can be used in the GGA-1/2 calculation with different GGA XCs, we have manually listed PBE and PBEsol as the baseline XCs. This means that we subsequently carried out GGA-1/2 calculations with all five XCs, but utilizing the set of self-energy potentials only from the baseline XCs. For all oxides, the corresponding band gap results are shown in table 5. Compared with table 2, the maximum GGA-1/2 band gap variation upon switching to the PBE-derived oxygen self-energy potential, is merely 0.02 eV among all XCs and all oxides. When exclusively using PBEsol-derived oxygen self-energy potential, the maximum band gap variation is enlarged to 0.06 eV, which is still a tiny value. We conclude that for these materials the self-energy correction is not significantly dependent on the XC flavor, and the PBE self-energy potentials may fit other XC flavors in carrying out GGA-1/2 calculations.

3.6. Influence of the self-energy potential trimming scheme

In the GGA-1/2 calculations, a cutoff function Θ is employed to trim the self-energy potential:

$$\Theta(r) = \begin{cases} \left[1 - \left(\dfrac{r}{r_{\text{cut}}}\right)^{p}\right]^{3}, & r \leqslant r_{\text{cut}} \\ 0, & r > r_{\text{cut}} \end{cases} \quad (6)$$

where p is a power index that should be sufficiently large to ensure a sharp trim, and r_{cut} is the cutoff radius. In the original work of Ferreira *et al* the value $p = 8$ was chosen to avoid possible convergence issues, but higher values can also be used. Indeed, we re-calculated the band gaps for the four oxides using GGA-1/2 with $p = 20$, and encountered no issue related to numerical convergence. Compared with $p = 8$, the GGA-1/2 band gaps with $p = 20$ (listed in table 6) increase by about 0.2 eV (for α-Al$_2$O$_3$) and 0.09 eV $\sim$ 0.14 eV (for r-TiO$_2$, m-ZrO$_2$ and m-HfO$_2$). The mean differences between the calculated band gap values and the experimental results are diminished as p increases from 8 to 20, indicating that $p = 20$ is more favorable for these calculations because it includes the self-energy more accurately. In order to explain this phenomenon, in figure 5 we plot the real-space spatial distribution of the hole (residing at the VBM in k-space) and the electron (residing at the CBM in k-space) for r-TiO$_2$ as an example. Since the hole and electron overlap, when the p value is low the self-energy potential reaches the electron density, lowering the CBM as well as the VBM. On the other hand, as p increases from 8 to 20, the self-energy potential acquires a sharper boundary, resulting in less overlap with the electron distribution, which yields larger band gaps. In the meantime, we notice a variation of the optimal oxygen self-energy cutoff radius when using different p values. As plotted in figure 6, the cutoff radii are in general reduced by around 0.2 a.u. $\sim$ 0.3 a.u. when p is increased from 8 to 20. Such phenomenon is common for all the materials under investigation. However, as long as p is fixed, the optimal r_{cut} is relatively constant across these oxides, showing good transferability of the self-energy potentials.

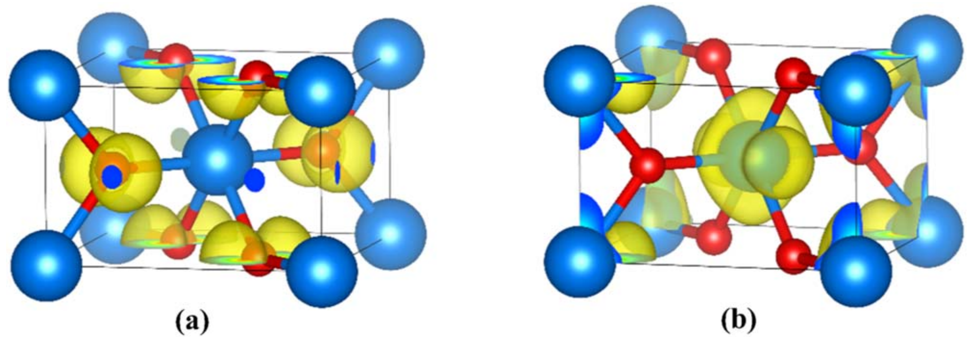

Figure 5. Spatial locations (yellow semi-transparent contours) for (a) the VB hole and (b) the CB electron in r-TiO$_2$. Big blue balls represent Ti while small red balls represent O.

Table 6. Calculated GGA-1/2 band gaps (unit: eV) for α-Al$_2$O$_3$, r-TiO$_2$, m-ZrO$_2$ and m-HfO$_2$ with different exchange-correlation functionals, with a power index $p = 20$ in the self-energy potential trimming function. The differences between band gaps obtained with $p = 8$ and $p = 20$ (ΔE_g) are also listed.

	PBE	PBEsol	PW91	revPBE	AM05	Expt.
			α-Al$_2$O$_3$			
$p = 20$	8.30	8.45	8.36	8.25	8.59	8.8 [53]
ΔE_g	0.20	0.20	0.20	0.21	0.21	
			r-TiO$_2$			
$p = 20$	3.13	3.18	3.06	3.11	3.18	3.0–3.4
ΔE_g	0.09	0.08	0.09	0.09	0.09	[58–61]
			m-ZrO$_2$			
$p = 20$	5.63	5.57	5.46	5.73	5.62	5.0–5.8 [62]
ΔE_g	0.12	0.12	0.13	0.14	0.12	5.4 [64]
			m-HfO$_2$			
$p = 20$	6.20	6.10	6.18	6.30	6.16	5.9 [67, 68]
ΔE_g	0.14	0.14	0.14	0.13	0.14	

3.7. Overall estimation of GGA-1/2

In figure 7 we compare the accuracy of the GGA-1/2 self-energy correction scheme under various XC flavors with conventional GGA, HSE06, and *GW*. Experimental values are used for benchmark. As expected, GGA band gaps are significantly underestimated compared to experimental values, differing by as much as 2.67 eV in the case of α-Al$_2$O$_3$ (see figure 7(a)). For the materials under investigation, the *GW* band gaps are the largest, except for m-HfO$_2$ where *GW* and GGA-1/2 are very close. Note that for m-HfO$_2$ the calculated *GW* band gaps vary considerably across the literature. For instance, in the case of m-HfO$_2$ Balá *et al* reported 5.45 eV band gap [62] while Jain *et al* reported 6.0 eV [66]. Moreover, for α-Al$_2$O$_3$ the *GW* band gaps are larger than experimental data by 0.56 eV $\sim$ 0.98 eV, while for r-TiO$_2$ and m-ZrO$_2$ the reported *GW* band gaps are in good agreement with experiments (see figures 7(b)–(d)).

The hybrid functional HSE06, depicted as hollow triangles in figure 7, clearly shows great improvement over conventional GGA, but its band gaps are generally smaller than those from *GW* calculations. On the other hand, the GGA-1/2 band gaps are generally better than HSE06 results, except for r-TiO$_2$. Among the five XC flavors, PBEsol demonstrates the best performance in general, in accordance with our previous results discussed in Section 3.3.

3.8. GGA-1/2 in the study of RRAM devices

After demonstrating the advantages of GGA-1/2 and identifying PBEsol as the XC with the best accuracy, we here list some possible applications of the GGA-1/2 method in investigating some fundamental problems in RRAM, especially regarding the nature of the CF:

- Prediction of CF composition and structure i
  n binary oxide-based RRAM cells. It is commonly agreed that oxygen vacancies (V_o's) are key to CF formation. However, for each material it is far from certain what amount of V_o should render CF formation with the

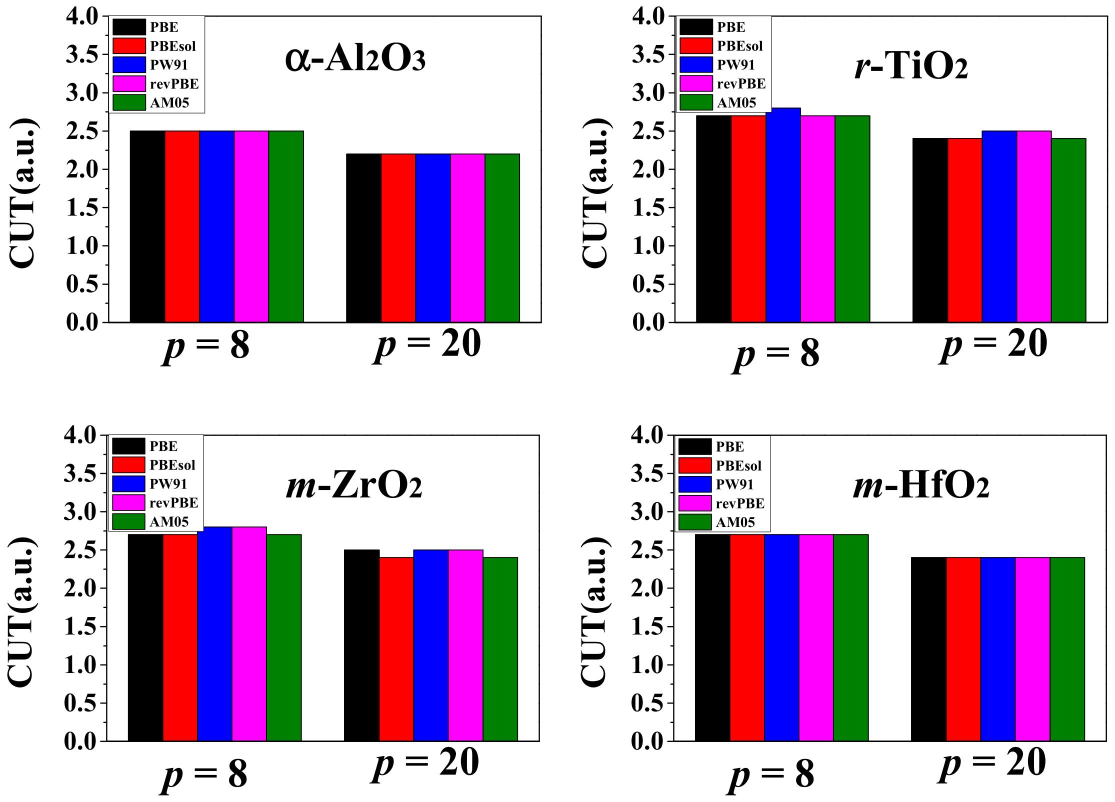

Figure 6. Optimal r_{cut} comparison for α-Al_2O_3, r-TiO_2, m-ZrO_2 and m-HfO_2 with different XC flavors, where the power index p is either 8 or 20.

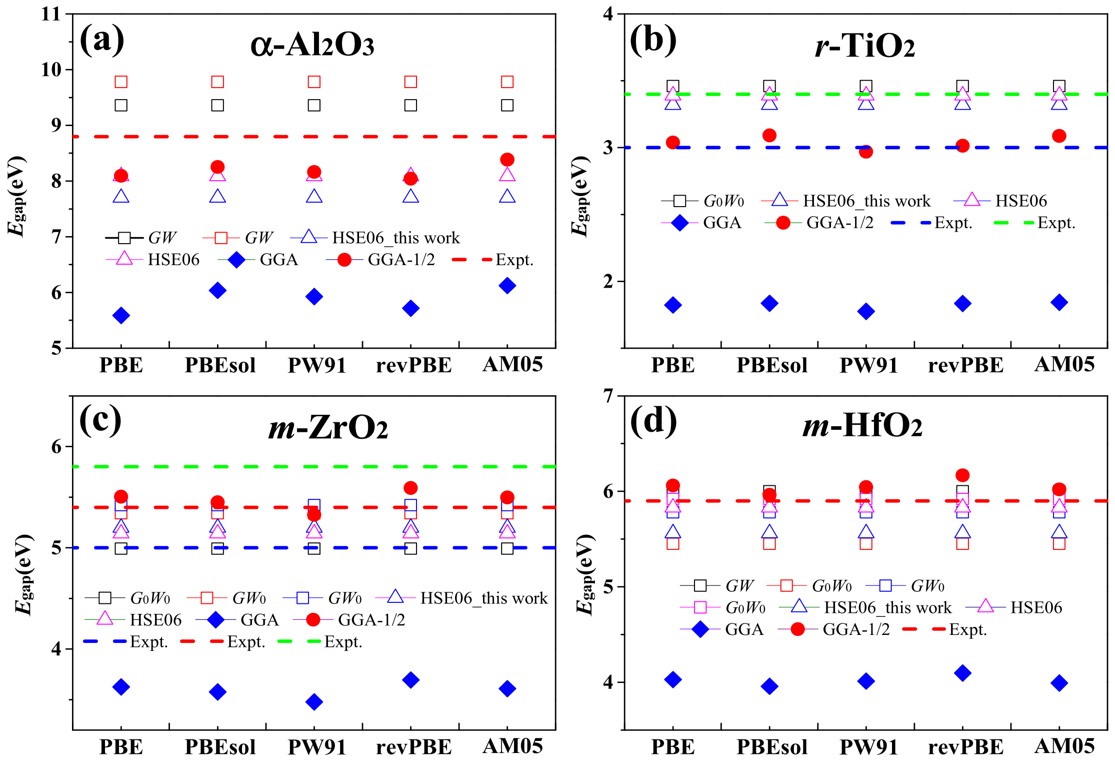

Figure 7. Energy band gaps for α-Al_2O_3, r-TiO_2, m-ZrO_2 and m-HfO_2 calculated with GGA (solid rhombuses), HSE06 (hollow triangles), GW (hollow squares), and GGA-1/2 (solid circles). Experimental values (horizontal dashed lines) are shown for comparison.

typical resistive switching behaviors found in experiments [70]. On the one hand, in the study of crystalline-phase sub-oxides [7], their conductive nature in principle ought to be predicted only by higher level techniques such as hybrid functionals and *GW* [24], since LDA/GGA incorrectly predicts the band gap, which tends to produce incorrect I-V characteristics. LDA/GGA can even predict a semiconductor to be a metal as it does to Ge [71, 72]. On the other hand, to evaluate more complex structures, such as amorphous phase CF candidates with various V_o concentrations, or CFs located in grain boundaries [73, 74], hybrid functionals and *GW* are computationally too demanding to be used. As we have shown, GGA-1/2 can efficiently deal with bulk crystalline oxides, as well as their amorphous phases or various interfaces.

- Investigation of the horizontal scaling limit of RRAM cells
  . The horizontal scaling limit means the minimum capacitor area that can sustain reliable LRS and HRS for a given RRAM technology. To this end one needs to establish the minimum stable area of a CF, and the minimum area of the insulating oxide phase surrounding the CF. The capability of maintaining the conductive state is important for the CF in small area capacitors, while the thermodynamic stability of the CF is another key factor. Hence, a CF-inside-oxide composite supercell should be studied using *ab initio* techniques. Initially we have used GGA to study the size limit of Hf suboxides inside HfO_2, even though the band gap was inaccurate [13]. Recently we have also investigated the area limit of Hf CFs inside HfO_2, where GGA-1/2 was implemented such that the band gap of HfO_2 was correctly recovered [75].
- Investigation of the vertical scaling limit of RRAM cells
  . In the vertical direction the RRAM cells also possess minimum thickness limit for the dielectric, since an even thinner dielectric may suffer from electron tunneling such that the HRS becomes leaky, which goes against the goal of a non-volatile memory. This requires setting up a 5 nm thick dielectric layer typically, encapsulated by two electrodes. Such setup can contain more than 500 atoms. Moreover, a series of dielectric thicknesses should be considered to extract the minimum dielectric thickness from the electronic structures of these capacitors. This computational task, which requires both band gap accuracy and numerical efficiency, can be addressed by GGA-1/2.
- Understanding device performance and reliability.

GGA-1/2 can also be used to investigate the reasons why a specific RRAM shows certain performance limitations. A typical example is the resistance window of the TiO_2 and HfO_2 RRAMs, where the latter is substantially larger [13, 76]. Wu *et al* [77] calculated the Schottky barriers of Ti_4O_7/TiO_2 using GGA-1/2, where Ti_4O_7 is the typical CF composition in TiO_2-based RRAMs [7]. They showed that the contact is Ohmic, thus the HRS of TiO_2 RRAMs does not possess the necessary barrier for electron injection from the CF to the dielectric. On the other hand, the CF/HfO_2 Schottky barriers are sufficiently large, no matter whether pure Hf or conductive sub-oxides were used as the CF model. GGA-1/2 is very useful for calculating the electronic structures of these models since previous calculations either suffered from band gap inaccuracy [78], or were forced to adopt short models for the dielectric to enable heavy hybrid functional calculations [31].

4. Conclusions

In summary, we performed a thorough investigation into the accuracy of the GGA-1/2 electronic structure calculation method, by comparing five different exchange-correlation functionals (PBE, PBEsol, PW91, revPBE and AM05). Four metal oxide crystals (α-Al_2O_3, *r*-TiO_2, *m*-ZrO_2 and *m*-HfO_2) were selected for benchmark calculations, based on their relevance in resistive memory devices, for which the simulation requires exceptionally large supercells in order to clarify the device physics. Such tasks ask for fast and accurate schemes for the calculation of band gaps and band offsets. We compared the accuracy of each GGA functional flavor and have drawn the following conclusions:

(1) GGA-1/2 within any of these exchange-correlation functional flavors predicts reasonable band gaps, correcting the band gap underestimation of conventional GGA.

(2) For the oxides investigated, the GGA-1/2 band gaps fall between the HSE06 and *GW* results, while the GGA-1/2 band structure morphology is in agreement with the computationally-demanding *GW* method.

(3) Since PBEsol and AM05 predict the most accurate lattice parameters, GGA-1/2 based on these two flavors yields band gaps that are relatively insensitive to the choice of experimental or optimized lattice parameters. In particular, PBEsol performs slightly better than AM05.

(4) While the atomic potentials obtained with different GGA flavors are fairly different, the corresponding self-energy potentials are relatively insensitive to the XC flavor that was used in the self-energy potential generation. In case a self-energy potential with a specified GGA XC is missing, one may resort to the PBE self-energy potential.

(5) In the implementation of GGA-1/2, increasing the power index of the self-energy cutoff function from 8 to 20 can further improve the calculated band gaps for typical metal oxides.

Our work shows that GGA-1/2 is an efficient and accurate method for calculating the electronic band structures for various resistive switching oxides, for all the GGA exchange-correlation functionals investigated. Nevertheless, the combination GGA(PBEsol)-1/2 stands out as the optimum scheme that offers the most accurate band gaps for these oxides.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (Materials Genome Initiative, 2017YFB0701700), the National Natural Science Foundation of China under Grant No. 11704134, the Fundamental Research Funds of Wuhan City under Grant No. 2017010201010106, the Fundamental Research Funds for the Central Universities of China under Grant No. HUST:2016YXMS212, and the Hubei "Chu-Tian Young Scholar" program. L R C Fonseca thanks the Brazilian agency CNPq for financial support under Grant No. 118485/2017-2.

ORCID iDs

Kan-Hao Xue https://orcid.org/0000-0002-2894-7912

References

[1] Walker S J 2014 Big data: a revolution that will transform how we live, work, and think *Int. J. Advert.* **33** 181–3
[2] Armbrust M *et al* 2010 A view of cloud computing *Commun. ACM* **53** 50-8
[3] Gubbi J, Buyya R, Marusic S and Palaniswami M 2013 Internet of things (IoT): a vision, architectural elements, and future directions *Future Gener. Comput. Syst.* **29** 1645–60
[4] Bez R, Camerlenghi E, Modelli A and Visconti A 2003 Introduction to flash memory *Proc. IEEE* **91** 489–502
[5] Pavan P, Bez R, Olivo P and Zanoni E 1997 Flash memory cells-an overview *Proc. IEEE* **85** 1248–71
[6] Gao B *et al* 2011 Oxide-based RRAM: Unified microscopic principle for both unipolar and bipolar switching *2011 Int. Electron Devices Meet.* pp 1–17
[7] Kwon D-H *et al* 2010 Atomic structure of conducting nanofilaments in TiO_2 resistive switching memory *Nat. Nanotechnol.* **5** 148–53
[8] Abbas Y, Sokolov A S, Jeon Y-R, Kim S, Ku B and Choi C 2018 Structural engineering of tantalum oxide based memristor and its electrical switching responses using rapid thermal annealing *J. Alloys Compd.* **759** 44–51
[9] Abbas Y, Jeon Y-R, Sokolov A S, Kim S, Ku B and Choi C 2018 Compliance-free, digital SET and analog RESET synaptic characteristics of sub-tantalum oxide based neuromorphic device *Sci. Rep.* **8** 1228
[10] Li Y, Zhong Y P, Zhang J J, Xu X H, Wang Q, Xu L, Sun H J and Miao X S 2013 Intrinsic memristance mechanism of crystalline stoichiometric $Ge_2Sb_2Te_5$ *Appl. Phys. Lett.* **103** 043501
[11] Liu X, Ji Z, Tu D, Shang L, Liu J, Liu M and Xie C 2009 Organic nonpolar nonvolatile resistive switching in poly(3,4-ethylene-dioxythiophene): Polystyrenesulfonate thin film *Org. Electron.* **10** 1191–4
[12] Wu Y, Lee B and Wong H S P 2010 Al_2O_3-based RRAM using atomic layer deposition (ALD) with 1-μA RESET current *IEEE Electron Device Lett.* **31** 1449–51
[13] Xue K-H, Traore B, Blaise P, Fonseca L R C, Vianello E, Molas G, De Salvo B, Ghibaudo G, Magyari-Kope B and Nishi Y 2014 A combined *ab initio* and experimental study on the nature of conductive filaments in Pt/HfO_2/Pt resistive random access memory *IEEE Trans. Electron Devices* **61** 1394–402
[14] Ku B, Abbas Y, Sokolov A S and Choi C 2018 Interface engineering of ALD HfO_2-based RRAM with Ar plasma treatment for reliable and uniform switching behaviors *J. Alloys Compd.* **735** 1181–8
[15] Zhao L, Zhang J, He Y, Guan X, Qian H and Yu Z 2011 Dynamic modeling and atomistic simulations of SET and RESET operations in TiO_2-based unipolar resistive memory *IEEE Electron Device Lett.* **32** 677–9
[16] Abbas H, Abbas Y, Truong S N, Min K-S, Park M R, Cho J, Yoon T-S and Kang C J 2017 A memristor crossbar array of titanium oxide for non-volatile memory and neuromorphic applications *Semicond. Sci. Technol.* **32** 065014
[17] Wu M-C, Lin Y-W, Jang W-Y, Lin C-H and Tseng T-Y 2011 Low-power and highly reliable multilevel operation in ZrO_2 1T1R RRAM *IEEE Electron Device Lett.* **32** 1026–8
[18] Guo Y and Robertson J 2015 *Ab initio* calculations of materials selection of oxides for resistive random access memories *Microelectron. Eng.* **147** 339–43
[19] Celano U, Goux L, Degraeve R, Fantini A, Richard O, Bender H, Jurczak M and Vandervorst W 2015 Imaging the three-dimensional conductive channel in filamentary-based oxide resistive switching memory *Nano Lett.* **15** 7970–5
[20] Celano U, Fantini A, Degraeve R, Jurczak M, Goux L and Vandervorst W 2016 Scalability of valence change memory: from devices to tip-induced filaments *AIP Adv.* **6** 085009
[21] Yang Y, Zhang X, Qin L, Zeng Q, Qiu X and Huang R 2017 Probing nanoscale oxygen ion motion in memristive systems *Nat. Commun.* **8** 15173

[22] Li C *et al* 2017 Direct observations of nanofilament evolution in switching processes in HfO_2-based resistive random access memory by *in situ* TEM studies *Adv. Mater.* **29** 1602976

[23] Kohn W and Sham L J 1965 Self-consistent equations including exchange and correlation effects *Phys. Rev.* **140** A1133–8

[24] Xue K-H, Blaise P, Fonseca L R C and Nishi Y 2013 Prediction of semimetallic tetragonal Hf_2O_3 and Zr_2O_3 from first principles *Phys. Rev. Lett.* **110** 065502

[25] Hedin L 1965 New method for calculating the one-particle green's function with application to the electron-gas problem *Phys. Rev.* **139** A796–823

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

[27] Park S-G, Magyari-Köpe B and Nishi Y 2010 Electronic correlation effects in reduced rutile TiO_2 within the LDA+U method *Phys. Rev.* B **82** 115109

[28] Park S-G, Magyari-Köpe B and Nishi Y 2009 Lattice and electronic effects in rutile TiO_2 containing charged oxygen defects from *ab initio* calculations *MRS Online Proc. Libr. Arch.* **1160** 1160-H11-11

[29] Park S G, Magyari-Kope B and Nishi Y 2008 First-principles study of resistance switching in rutile TiO_2 with oxygen vacancy *9th Annu. Non-Volatile Mem. Technol. Symp. NVMTS* pp 1–5

[30] Zhao L, Park S-G, Magyari-Köpe B and Nishi Y 2013 Dopant selection rules for desired electronic structure and vacancy formation characteristics of TiO_2 resistive memory *Appl. Phys. Lett.* **102** 083506

[31] Li H, Zhang Z and Shi L P 2016 Identifying and engineering the electronic properties of the resistive switching interface *J. Electron. Mater.* **45** 1142–53

[32] Ferreira L G, Marques M and Teles L K 2008 Approximation to density functional theory for the calculation of band gaps of semiconductors *Phys. Rev.* B **78** 125116

[33] Ferreira L G, Marques M and Teles L K 2011 Slater half-occupation technique revisited: the LDA-1/2 and GGA-1/2 approaches for atomic ionization energies and band gaps in semiconductors *AIP Adv.* **1** 032119

[34] Perdew J P, Ruzsinszky A, Csonka G I, Vydrov O A, Scuseria G E, Constantin L A, Zhou X and Burke K 2008 Restoring the density-gradient expansion for exchange in solids and surfaces *Phys. Rev. Lett.* **100** 136406

[35] Wu Z and Cohen R E 2006 More accurate generalized gradient approximation for solids *Phys. Rev.* B **73** 235116

[36] Armiento R and Mattsson A E 2005 Functional designed to include surface effects in self-consistent density functional theory *Phys. Rev.* B **72** 085108

[37] Levin I and Brandon D 1998 Metastable alumina polymorphs: Crystal structures and transition sequences *J. Am. Ceram. Soc.* **81** 1995–2012

[38] Abrahams S C and Bernstein J L 1971 Rutile: Normal probability plot analysis and accurate measurement of crystal structure *J. Chem. Phys.* **55** 3206–11

[39] Miikkulainen V, Leskelä M, Ritala M and Puurunen R L 2013 Crystallinity of inorganic films grown by atomic layer deposition: Overview and general trends *J. Appl. Phys.* **113** 021301

[40] Kresse G and Furthmüller J 1996 Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set *Phys. Rev.* B **54** 11169–86

[41] Kresse G and Furthmüller J 1996 Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set *Comput. Mater. Sci.* **6** 15–50

[42] Perdew J P, Burke K and Ernzerhof M 1996 Generalized gradient approximation made simple *Phys. Rev. Lett.* **77** 3865–8

[43] Perdew J P, Chevary J A, Vosko S H, Jackson K A, Pederson M R, Singh D J and Fiolhais C 1992 Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation *Phys. Rev.* B **46** 6671–87

[44] Hammer B, Hansen L B and Nørskov J K 1999 Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals *Phys. Rev.* B **59** 7413–21

[45] Blöchl P E 1994 Projector augmented-wave method *Phys. Rev.* B **50** 17953–79

[46] Kresse G and Joubert D 1999 From ultrasoft pseudopotentials to the projector augmented-wave method *Phys. Rev.* B **59** 1758–75

[47] Monkhorst H J and Pack J D 1976 Special points for Brillouin-zone integrations *Phys. Rev.* B **13** 5188–92

[48] Leite J R and Ferreira L G 1971 Effects of the coulomb correlation on the calculated results for atoms with and without spin polarization *Phys. Rev.* A **3** 1224–30

[49] Soler J M, Artacho E, Gale J D, García A, Junquera J, Ordejón P and Sánchez-Portal D 2002 The SIESTA method for *ab initio* order-N materials simulation *J. Phys. Condens. Matter* **14** 2745–79

[50] Oetzel M and Heger G 1999 Laboratory x-ray powder diffraction: a comparison of different geometries with special attention to the usage of the Cu $K\alpha$ doublet *J. Appl. Crystallogr.* **32** 799–807

[51] Murugesan S, Kuppusami P and Mohandas E 2010 Rietveld X-ray diffraction analysis of nanostructured rutile films of titania prepared by pulsed laser deposition *Mater. Res. Bull.* **45** 6–9

[52] Whittle K R, Lumpkin G R and Ashbrook S E 2006 Neutron diffraction and MAS NMR of cesium tungstate defect pyrochlores *J. Solid State Chem.* **179** 512–21

[53] French R H, Jones D J and Loughin S 1994 Interband electronic structure of α-alumina up to 2167 K *J. Am. Ceram. Soc.* **77** 412–22

[54] Santos R C R, Longhinotti E, Freire V N, Reimberg R B and Caetano E W S 2015 Elucidating the high-k insulator α-Al_2O_3 direct/indirect energy band gap type through density functional theory computations *Chem. Phys. Lett.* **637** 172–6

[55] Marinopoulos A G and Grüning M 2011 Local-field and excitonic effects in the optical response of α-alumina *Phys. Rev.* B **83** 195129

[56] Peng H, Scanlon D O, Stevanovic V, Vidal J, Watson G W and Lany S 2013 Convergence of density and hybrid functional defect calculations for compound semiconductors *Phys. Rev.* B **88** 115201

[57] Landmann M, Rauls E and Schmidt W G 2012 The electronic structure and optical response of rutile, anatase and brookite TiO_2 *J. Phys. Condens. Matter* **24** 195503

[58] Burdett J K, Hughbanks T, Miller G J, Richardson Jr J W and Smith J V 1987 Structural-electronic relationships in inorganic solids: powder neutron diffraction studies of the rutile and anatase polymorphs of titanium dioxide at 15 and 295 K *J. Am. Chem. Soc.* **109** 3639–46

[59] Amtout A and Leonelli R 1995 Optical properties of rutile near its fundamental band gap *Phys. Rev.* B **51** 6842–51

[60] Clark S J and Robertson J 2010 Screened exchange density functional applied to solids *Phys. Rev.* B **82** 085208

[61] Pascual J, Camassel J and Mathieu H 1977 Resolved quadrupolar transition in TiO_2 *Phys. Rev. Lett.* **39** 1490–3

[62] Jiang H, Gomez-Abal R I, Rinke P and Scheffler M 2010 Electronic band structure of zirconia and hafnia polymorphs from the *GW* perspective *Phys. Rev.* B **81** 085119

[63] Sinhamahapatra A, Jeon J-P, Kang J, Han B and Yu J-S 2016 Oxygen-deficient zirconia (ZrO_{2-x}): A new material for solar light absorption *Sci. Rep.* **6** 27218

[64] Králik B, Chang E K and Louie S G 1998 Structural properties and quasiparticle band structure of zirconia *Phys. Rev.* B **57** 7027–36

[65] Houssa M, Tuominen M, Naili M, Afanas'ev V, Stesmans A, Haukka S and Heyns M M 2000 Trap-assisted tunneling in high permittivity gate dielectric stacks *J. Appl. Phys.* **87** 8615–20

[66] Jain M, Chelikowsky J R and Louie S G 2011 Quasiparticle excitations and charge transition levels of oxygen vacancies in hafnia *Phys. Rev. Lett.* **107** 216803

[67] Sayan S, Emge T, Garfunkel E, Zhao X, Wielunski L, Bartynski R A, Vanderbilt D, Suehle J S, Suzer S and Banaszak-Holl M 2004 Band alignment issues related to HfO_2/SiO_2/*p*-Si gate stacks *J. Appl. Phys.* **96** 7485–91

[68] Alkauskas A, Broqvist P and Pasquarello A 2008 Defect energy levels in density functional calculations: Alignment and band gap problem *Phys. Rev. Lett.* **101** 046405

[69] Chen W and Pasquarello A 2012 Band-edge levels in semiconductors and insulators: Hybrid density functional theory versus many-body perturbation theory *Phys. Rev. B* **86** 035134

[70] Kim K M, Jeong D S and Hwang C S 2011 Nanofilamentary resistive switching in binary oxide system; a review on the present status and outlook *Nanotechnology* **22** 254002

[71] Broqvist P, Alkauskas A and Pasquarello A 2008 Defect levels of dangling bonds in silicon and germanium through hybrid functionals *Phys. Rev.* B **78** 075203

[72] Xue K-H, Yuan J-H, Fonseca L R C and Miao X-S 2018 Improved LDA-1/2 method for band structure calculations in covalent semiconductors *Comput. Mater. Sci.* **153** 493–505

[73] Xue K-H, Blaise P, Fonseca L R C, Molas G, Vianello E, Traoré B, De Salvo B, Ghibaudo G and Nishi Y 2013 Grain boundary composition and conduction in HfO_2: An *ab initio* study *Appl. Phys. Lett.* **102** 201908

[74] Shubhakar K, Mei S, Bosman M, Raghavan N, Ranjan A, O'Shea S J and Pey K L 2016 Conductive filament formation at grain boundary locations in polycrystalline HfO_2-based MIM stacks: Computational and physical insight *Microelectron. Reliab.* **64** 204–9

[75] Xue K-H and Miao X-S 2018 Oxygen vacancy chain and conductive filament formation in hafnia *J. Appl. Phys.* **123** 161505

[76] Biju K P, Liu X J, Bourim E M, Kim I, Jung S, Siddik M, Lee J and Hwang H 2010 Asymmetric bipolar resistive switching in solution-processed Pt/TiO_2/W devices *J. Phys. D: Appl. Phys.* **43** 495104

[77] Wu Z-H, Xue K-H and Miao X-S 2017 Filament-to-dielectric band alignments in TiO_2 and HfO_2 resistive RAMs *J. Comput. Electron.* **16** 1057–65

[78] Gu T, Wang Z, Tada T and Watanabe S 2009 First-principles simulations on bulk Ta_2O_5 and Cu/Ta_2O_5/Pt heterojunction: Electronic structures and transport properties *J. Appl. Phys.* **106** 103713