# Anomalous behavior of the semiconducting gap in $WO_3$ from first-principles calculations

G. A. de Wijs, P. K. de Boer, and R. A. de Groot
*Electronic Structure of Materials, Research Institute of Materials, Faculty of Sciences, Toernooiveld 1, NL-6525 ED Nijmegen, The Netherlands*

G. Kresse
*Institute for Theoretical Physics, Technical University of Vienna, Wiedner Hauptstraße 8-10/136, A-1040 Vienna, Austria*
(Received 13 May 1998; revised manuscript received 11 September 1998)

Several crystal structures of tungsten trioxide have been studied with a first-principles pseudopotential method. The electronic band gap increases significantly with the distortion of the octahedra that are the building blocks of the various crystal structures. Moreover, the tilting of the octahedra in the more complex structures leads to a strong increase of the gap upon compression. [S0163-1829(99)04104-1]

## I. INTRODUCTION

Tungsten trioxide has received much attention from the experimental point of view because of its potential for technological applications. Ion intercalation and deintercalation into its open structure, going hand in hand with insertion and extraction of electrons, e.g., by application of an electric field, cause a strong modulation of its optical properties (electrochromism). For an exhaustive review on these so-called tungsten bronzes see Ref. 1.

Naturally, this material has also attracted the attention of computational physics and many useful studies have appeared in the literature so far. Non-self-consistent and/or semiempirical calculations were reported by Kopp *et al.*,[2] Bullett,[3,4] Zhan and Zheng,[5] and Stashans and Lunell.[6] Self-consistent, first-principles studies were carried out by Christensen and Mackintosh,[7] Corà *et al.*,[8] Stachiotti *et al.*,[9] and Hjelm *et al.*[10] Except for the Hartree-Fock studies by Corà *et al.* these were all density-functional-based calculations.

The first-principles studies have only considered the simplified, i.e., simple cubic, perovskitelike structure as depicted in Fig. 1, with the exception of the model hexagonal structure studied in Ref. 10 and the tetragonal structure in Ref. 8. The simplified perovskitelike structure indeed occurs for high fractions of intercalated material (e.g., for $NaWO_3$), but it is not more than a first approximation to the structures of the pure $WO_3$ single crystals. It exhibits the basic structural characteristics, i.e., corner-sharing oxygen octahedra enclosing the tungsten atoms, but ignores important details concerning their distorted shape and their relative orientation. In fact, depending on these details, five different phases are known as a function of temperature: from (approximately) $-140\,^{\circ}\mathrm{C}$ to $-50\,^{\circ}\mathrm{C}$ a monoclinic low-temperature (LT) phase occurs, from $-50\,^{\circ}\mathrm{C}$ to $17\,^{\circ}\mathrm{C}$ a triclinic phase exists, followed by a monoclinic room-temperature (RT) modification that is stable to about $330\,^{\circ}\mathrm{C}$. Above that temperature, and until $740\,^{\circ}\mathrm{C}$, $WO_3$ becomes orthorhombic, and finally, above that temperature a tetragonal phase has been found.[11] Studies on some of these crystals have been carried out with the non-self-consistent methods (Refs. 3–6). Those by Bullett have demonstrated that in the simplified cubic structure the semiconducting gap is significantly underestimated compared to the monoclinic RT structure.

In this paper we investigate several of the $WO_3$ structures occurring in reality by means of a self-consistent first-principles method carrying out geometry optimizations. We demonstrate that the electronic structure strongly correlates with the distortions of the octahedra and their mutual orientation (thereby confirming the non-self-consistent results by Bullett) and analyze this behavior. Further we discuss how the electronic structure changes upon compression and expansion and relate this with the deformation mechanisms occurring. Section II presents details on the method of calculation. In Sec. III results of the calculations on various structures are presented. Conclusions and discussion can be found in Sec. IV.

## II. CALCULATIONAL APPROACH

The calculations have been performed using the *ab initio* total-energy and molecular-dynamics program VASP (Vienna *Ab initio* Simulation Program) developed at the Institut für Theoretische Physik of the Technische Universität Wien.[12–14]

Electron-ion interactions were described using Vanderbilt-type (Ref. 15) ultrasoft pseudopotentials (USPP) with a frozen $[\mathrm{Xe}]4f^{14}$ and $1s^2$ core for W and O, respectively, as supplied by the Institut für Theoretische Physik.[16] Nonlinear core corrections were applied for W.[17] For efficiency, in calculations on larger cells, the real-space

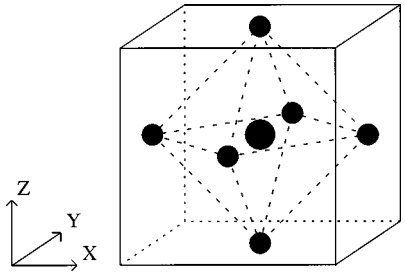


FIG. 1. Cubic perovskitelike crystal structure. The small spheres at the centers of the cubes faces represent oxygen atoms, the large sphere in the center of the cube represents a tungsten atom.

 

TABLE I. Convergence test results for simple cubic $WO_3$ in the GGA. Equilibrium volume $V_{eq}$, bulk modulus $B$, and total energy $E$ as a function of cutoff $E_{PW}$ and **k**-point mesh.

| $E_{PW}$ (eV) | Mesh | $V_{eq}$(Å$^3$) | $B$ (GPa) | $E$ (eV) |
|---|---|---|---|---|
| 400 | $4\times4\times4$ | 56.3487 | 224.1 | −36.267 |
| 400 | $6\times6\times6$ | 56.3251 | 224.2 | −36.227 |
| 400 | $8\times8\times8$ | 56.3369 | 223.2 | −36.222 |
| 500 | $6\times6\times6$ | 56.4308 | 226.2 | −36.220 |
| Real space projection | | | | |
| 400 | $6\times6\times6$ | 56.3128 | 224.5 | −36.232 |

projection scheme by King-Smith *et al.* for the nonlocal part of the USPP was used.[18]

Exchange and correlation were treated in the generalized gradient approximation (GGA) based on the parametrization by Perdew and Zunger (Ref. 19) of the local-density functional of Ceperley and Alder (Ref. 20) with the gradient corrections following Perdew *et al.* (PW 91).[21] Several calculations were redone using the local-density functional only (with pseudopotentials generated consistently). This was done in order to facilitate comparison with other calculations and to test whether it can cure the overestimation of the equilibrium volumes found with the GGA.

Several **k**-point meshes (Ref. 22) were used, where the density depended on the property that needed to be calculated.

Convergence tests were carried out on the primitive cell of the simple cubic structure (Fig. 1). Results are reported in Table I. These were obtained from a Murnaghan fit to a set of total-energy calculations at several volumes and fixed plane-wave cutoff $E_{PW}$ on the orbitals. From the results it follows that an $E_{PW}=400$ eV (30 Ry) is sufficient. A $4\times4\times4$ **k**-point mesh [i.e., 4 **k** points in the irreducible part of the cubic Brillouin zone (BZ)] is already sufficient to converge to the equilibrium structure, but for the total energy at least a $6\times6\times6$ mesh is desirable. When carrying out structural relaxations we therefore adopted a mesh consistent with (or better than) the $4\times4\times4$ mesh of this small cell. For example, in a monoclinic cell containing eight $WO_3$ units in a nearly cubic supercell we used a $2\times2\times2$ mesh. Total energies were calculated from the equilibrium structures thus obtained in extra calculations using a mesh at least as dense as, and preferably consistent with, the $6\times6\times6$ mesh of the test system. So for the monoclinic cell mentioned before we used a $3\times3\times3$ mesh shifted away from $\Gamma$.

Relaxations were always carried out at constant cutoff, i.e., the equilibrium volume is obtained from a fit to a set of calculations at various fixed volumes. Positional parameters (if any) have been relaxed for each volume separately.

Since the $WO_3$ structures studied in this paper have a gap of $\approx0.5$ eV or larger, the results are rather insensitive to the kind of Fermi-surface smearing employed. We have used both Gaussian smearing and the (modified) tetrahedron method.[23]

## III. RESULTS

Table II lists the total energies and equilibrium volumes of all crystal structures studied. Also shown are the results

TABLE II. Equilibrium volumes $V_{eq}$ and total energies $E$ for various crystal structures.

| | | $V_{eq}$ (Å$^3$/$WO_3$) | $E$(eV/$WO_3$) | $B$(GPa) |
|---|---|---|---|---|
| | | Simple cubic | | |
| | GGA | 56.3 | −36.23 | 224 |
| | LDA | 54.2 | −39.66 | 256 |
| FP-LMTO [a] | LDA | 56.6 | | |
| FP-LMTO [b] | LDA | 54.0 | | 254 |
| | | Tetragonal | | |
| | GGA | 57.2 | −36.322 | |
| | | Monoclinic, RT | | |
| | GGA | 56.3 | −36.341 | |
| | LDA | 52.3 | −39.752 | ∼40 |
| Expt. [c] | | 52.96 | | |
| | | Monoclinic, LT & HP | | |
| | GGA | 55.8 | −36.336 | |
| | LDA | 52.7 | −39.751 | |
| LT, expt. [d] | | 52.08 | | |
| | | Triclinic, LT | | |
| | GGA | 56.5 | −36.342 | |
| Expt. [e] | | 52.74 | | |
| | | Simple cubic $NaWO_3$ | | |
| | GGA | 57.0 | −39.84 | 203 |
| | LDA | 57.2 | −43.73 | 235 |
| FP-LMTO [a] | LDA | 58.4 | | |
| FP-LMTO [b] | LDA | 56.2 | | 241 |
| Expt. [f] | | 58.0 | | |

[a]Reference 10.
[b]Reference 9.
[c]Neutron-diffraction experiment, Ref. 25.
[d]Reference 37.
[e]Experiment at 10 °C, Ref. 35.
[f]Experiment, Refs. 11 and 36.

for the simple cubic structure by Hjelm *et al.*[10] and Stachiotti *et al.*[9] as obtained with the full-potential linear muffin-tin orbitals (FP-LMTO) method in the local-density approximation (LDA). These were obtained with other parametrizations of the LDA functional than our results based on the Ceperley-Alder functional. We do not expect this to affect the results, and FP-LMTO calculations should serve as a good benchmark for our calculations. Indeed our results, both lattice constant and bulk modulus, are very close to those of Ref. 9. Somewhat surprisingly, the equilibrium volume reported by Hjelm *et al.* is somewhat larger (which amounts to a lattice constant $\approx1.5\%$ larger). The GGA causes an increase of the equilibrium volume, in accordance with the general trend of GGA enhancing interatomic separations over LDA. For simple cubic $NaWO_3$ we obtain an equilibrium volume in between both FP-LMTO results.

Comparison to experiment is most conveniently done for the monoclinic structure for which accurate RT neutron-diffraction data are available.[24,25] Our GGA calculations overestimate the equilibrium volume (a discrepancy of 2.1% in the lattice constant) whereas the LDA calculations come very close to the experimental number. This is in agreement with the tendency of GGA to increase lattice constants compared to LDA, but—unlike the simple cubic structure—the

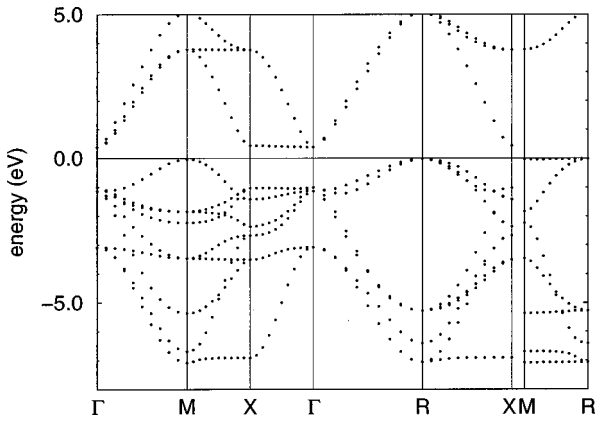


FIG. 2. GGA band structure of simple cubic $WO_3$ at $V = 56.3$ Å$^3$. The oxygen $2s$ bands are below the frame of the plot.

bond length elongation cannot account for the effect (this is discussed in detail in Sec. III D). Moreover, it is known that in perovskites GGA tends to overestimate the equilibrium volume. For $NaWO_3$ a similar overestimation of the lattice constant with the GGA occurs, but by only 1.1%.

The total energy of the artificial simple cubic tungsten trioxide is approximately 0.1 eV per formula unit above the total energies of the other structures. The latter only differ by a few meV among themselves. Entropic effects, which we expect to be of similar size, will decide what structure is most stable at which temperature. Only the tetragonal structure drops out, its total energy being about 0.02 eV above that of the others. However, this is the highest-temperature phase, occurring above 740 °C.

Below various structures are discussed in detail separately. It is shown how the electronic gap increases with increasing complexity of the deformation of the octahedra and how this affects the mechanism underlying the volume dependence of the gap.

### A. Simple cubic $WO_3$

$WO_3$ is a compound with a strong ionic character. In a simplified picture, all six W valence electrons are transferred to the O atoms, who can accommodate two additional electrons each, leading to the formation of an insulator at the stoichiometry $WO_3$. The band structure (Fig. 2) of simple cubic (sc) $WO_3$ (whose structure is depicted in Fig. 1) is in accordance with this qualitative reasoning: An oxygen-derived valence band is separated by a small gap from a tungsten-derived conduction band. The gap is indirect, occurring between the line $M$-$R$ in the valence and the line $X$-$\Gamma$ in the conduction region.

One of the states at the top of the valence band (located at $M$) is depicted schematically in Fig. 3(a). It is evident that all interactions between the $2p$ functions of nearest-neighbor oxygen atoms (in the $x$-$y$ plane) are of antibonding character, which explains why this state is at the top of the valence band. Moreover, the interactions of this state with any tungsten $5d$ function or $2p$ orbital on the other oxygen atoms are of nonbonding character. So all interactions are restricted to the $x$-$y$ plane, and dispersion is absent along the line $M$-$R$.

The conduction-band structure reflects the octahedral crystal field felt by the tungsten atoms; above the gap one

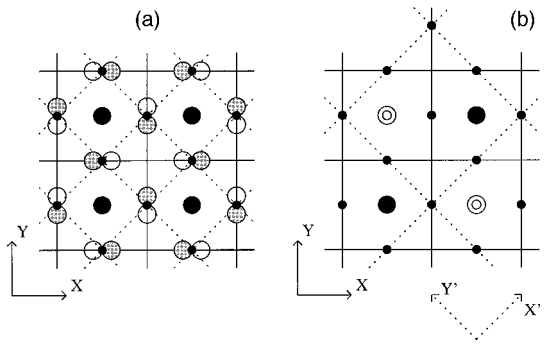


FIG. 3. (a) One of the states at the top of the valence band ($M$) in the simple cubic structure. Solid lines separate a sc unit cell from its periodic images. The dotted lines connect nearest-neighbor oxygen atoms in the $x$-$y$ plane. The W and O are represented by large and small bullets respectively. (b) Relation between the tetragonal and simple cubic structure. In the tetragonal structure the W represented by two concentric circles have moved into the positive $z$ direction, whereas those represented by the filled circles have moved in the opposite direction. The unit cell contains two formula units and is separated from its periodic images by the dashed lines. Note that it is rotated by 45° around the $z$ axis.

finds the $t_{2g}$ orbitals ($5d_{xy}$, $5d_{yz}$, and $5d_{xz}$ derived) whereas the $e_g$ orbitals are located at much higher energies. The conduction-band minimum occurs at $\Gamma$, and one of the bands is dispersionless along $\Gamma$-$X$ (reflecting the absence of interaction for, e.g., the $5d_{xy}$-derived band along $z$).

The band structure is in qualitative agreement with the results by Bullett,[3] the most striking difference being that Bullett finds the top of the valence band occurring at $\Gamma$. The magnitude of the gap (0.4 eV) agrees reasonably well with the 0.6 eV found with FP-LMTO in Ref. 10. The authors of Ref. 10 do not report a band structure, but the width of the valence part of their DOS closely agrees with the energy range spanned by the valence bands in Fig. 2 (both are close to 7 eV). In Ref. 9 (also FP-LMTO) a band structure is reported, whose shape is similar to ours, except for some small differences; e.g., a somewhat stronger dispersion near the Fermi level along the lines $M$-$R$ and $X$-$\Gamma$. However, the valence energy bandwidth is 1 eV larger as it is in our calculation. This discrepancy is too large to be explained by the different treatment of exchange and correlation (GGA vs LDA) or by its indirect consequences (a different volume). Indeed, it remains enigmatic, since equilibrium volume and bulk modulus from Ref. 9 and ours agree very closely. Moreover, this discrepancy also exist between Ref. 9 and Ref. 10, and both of these employ a very similar computational method. A band structure was also reported in Ref. 8. A detailed comparison is not very useful, since it is based on a rather different technique (Hartree-Fock), but results agree qualitatively.

The gap is weakly volume dependent and increases upon compression. This behavior is consistent with an ionic picture. Considering the limit of high compression, the kinetic-energy contribution will dominate the total energy, and higher kinetic-energy bands will move upwards on the energy scale more rapid under compression (see, e.g., Ref. 26). The simple cubic $WO_3$ will evolve towards this limit, and the $5d$-derived conduction bands (having two radial nodes and therefore higher kinetic energy) will move upwards with re-

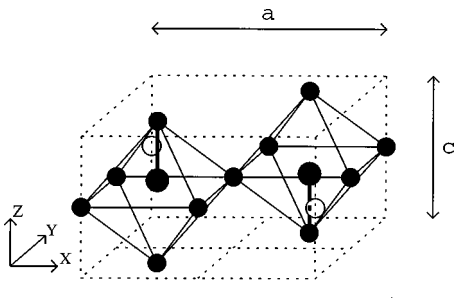


FIG. 4. Tetragonal structure. Large spheres and small spheres represent the tungsten and oxygen atoms, respectively. Oxygen atoms not partaking of the two octahedra fully shown are represented by open circles. Along *z* the short W-O bonds are represented by the fat lines.

spect to the oxygen $2p$-derived valence bands (having no radial nodes). Counteracting this tendency to increase the gap is the broadening of the bands under compression. The band broadening is expected to be stronger in a covalently bonded material, where the increasing overlap between nearest-neighbor-centered ''atomic orbitals'' is responsible for the broadening. On the contrary, in an ionic material, next-nearest-neighbor interactions determine the bandwidth, resulting in a weaker volume dependence. In the present case, the combined effect of these two mechanisms is an increase of the gap with compression, which we interpret as the material behaving in an essentially ionic fashion. Moreover, we note the absence of any interaction [notably no bonding-antibonding (BAB) splitting] between the top of the valence and the bottom of the conduction band.

### B. Tetragonal modification

The tetragonal structure is slightly more complicated than the simple cubic structure. It is depicted schematically in Fig. 4 and contains two formula units per unit cell. In the *x*-*y* plane all W-O bonds are of equal length, but in the *z* direction alternatingly long and short W-O bonds occur. This happens in an ''antiferroelectric'' fashion, since the two infinite W-O chains that pass through the unit cell (as depicted in Fig. 4) can be transformed into one another by an inversion through the center of this cell, i.e., they run in opposite directions. In addition to the relaxation of the $c/a$ ratio, in the calculations we have to allow for a relaxation along *z* of the central W atom and its oxygen neighbors in the infinite W-O chain running in the same direction (the relaxations in the other octahedron follow as a consequence of the inversion symmetry). The relation between the unit cell of the tetragonal modification and the simple cubic unit cell is shown schematically in Fig. 3(b). Note that the *x* and *y* coordinates are rotated by 45°.

The lattice dimensions, calculated within the GGA, are compared to experiment (Refs. 27 and 28) in Table III. We see a slight overestimation of *a* and *c*, just as for the other structures listed in the table (see Sec. III D). Due to limited accuracy of the original experiment the authors of Ref. 27 could not determine whether the long-short alternation of W-O bonds along the *z* axis (as found by our calculations, see below) occurs. However, their (high-temperature) study showed that the W atoms lie approximately 0.23 Å above the plane defined by their four nearest neighbors along *x* and *y*. We find 0.27 Å for this distance.

TABLE III. Lattice dimensions. ''Expt.'' denotes experimental data.

| | $a$ (Å) | $b$ (Å) | $c$ (Å) | $\alpha$ | $\beta$ | $\gamma$ |
|---|---|---|---|---|---|---|
| | | | Tetragonal | | | |
| Expt. [a] | 5.19 | | 3.86 | | | |
| Expt. [b] | 5.250 | | 3.915 | | | |
| Exp. [c] | 5.272 | | 3.920 | | | |
| GGA | 5.36 | | 3.98 | | | |
| | | | Monoclinic LT & HP | | | |
| LT, Expt. [d] | 5.277 | 5.155 | 7.662 | | 91.76° | |
| GGA | 5.37 | 5.31 | 7.81 | | 90.8° | |
| LDA | 5.34 | 5.31 | 7.77 | | 90.6° | |
| | | | Monoclinic RT | | | |
| Expt. [e] | 7.306 | 7.540 | 7.692 | | 90.88° | |
| GGA | 7.55 | 7.62 | 7.83 | | 90.2° | |
| LDA | 7.37 | 7.46 | 7.64 | | 90.6° | |
| | | | Triclinic | | | |
| Expt. [f] | 7.30 | 7.52 | 7.69 | 88.85° | 90.92° | 90.95° |
| GGA | 7.54 | 7.64 | 7.84 | 89.7° | 90.2° | 90.2° |

[a] $T = 740\,^{\circ}\mathrm{C}$, Ref. 27.
[b] $T = 770\,^{\circ}\mathrm{C}$, Ref. 28.
[c] $T = 950\,^{\circ}\mathrm{C}$, Ref. 28.
[d] $T = 5$ K, Ref. 37.
[e] Reference 25.
[f] Experiment at $10\,^{\circ}\mathrm{C}$, Ref. 35.

Using symmetry arguments only, one would expect that the gap be direct since one of the three points *M* of the simple cubic BZ now folds into Γ [e.g., the state in Fig. 3(a) folds back to Γ]. This guess is confirmed in the calculated band structure as shown in Fig. 5. However, the degeneracy of the three lowest bands above the gap at Γ is partly lifted: a group of two bands, strongly curved and degenerate along Γ-*Z*, and one isolated band, showing no interaction along Γ-*Z* (i.e., the direction corresponding to the *z* axis in real space) remain. The former bands lie more than twice farther away from the top of the valence band than the latter band.

The high shift upwards of this group of two bands at Γ can be understood on the basis of the alternatingly long and

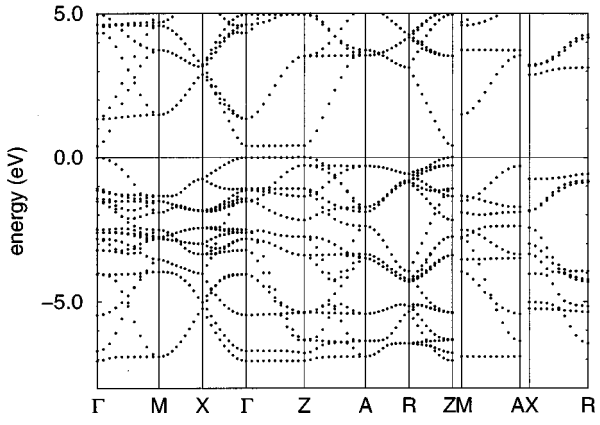


FIG. 5. GGA band structure of tetragonal $WO_3$ at $V = 57.2$ Å$^3$.

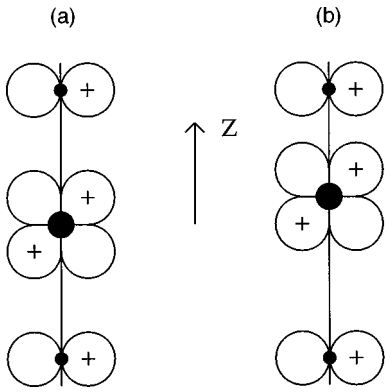


FIG. 6. Schematic representation of the interaction between the $d_{xz}$ and $p_x$ (or $d_{yz}$ and $p_y$) at $\Gamma$. (a) The simple cubic structure, (b) the tetragonal structure.

short W-O bonds in the $z$ direction. The consequences can be understood from Fig. 6, which depicts the relevant interactions at $\Gamma$. Whereas (because of the symmetry) in the sc structure the W $5d_{yz}$ and the $2p_y$ of a neighboring O atom (along $z$) cannot have any overlap at $\Gamma$ [Fig. 6(a)], the movement of the oxygen towards one of its tungsten neighbors allows for such an overlap to be developed [Fig. 6(b)]. This also holds for the $5d_{xz}$ and the oxygen $2p_x$. However, the two tungsten atoms in the unit cell move in opposite directions, so where the tungsten atom moving in the positive $z$ direction can develop a significant overlap with its O neighbor on top, the tungsten atom moving in the negative $z$ direction can only develop a significant overlap with its oxygen neighbor below. For any interaction to result, the overlaps should not cancel but have the same sign. Therefore, the phase of the oxygen functions (which, at $\Gamma$, is the same above and below a tungsten atom) will have to change by $\pi$ going from an O $2p_x$ or $2p_y$ centered above (below) one of the W to an O-centered function above (below) the other W. We have such oxygen states available now at $\Gamma$, as they are folded back from one of the $M$ points in the simple cubic BZ (they are situated $\sim 2$ eV below the Fermi level in the sc BZ, i.e., well below the valence-band top). This mechanism results in a nonzero overlap with the W $5d_{xz}$- and $5d_{yz}$-derived states (located at $\Gamma$ in *both* the sc and tetragonal structures). An interaction of BAB-type occurs, which results in a movement upwards of the $5d_{xz}$- and $5d_{yz}$-derived bands in the conduction region, as well as a downward movement of several bands in the valence region (this downward movement, resulting in a lowering of the band-structure energy, provides the driving force for the long-short splitting). However, this mechanism cannot occur for $5d_{xy}$, since all W-O bonds are of equal length in both the $x$ and $y$ directions [seen from the positive $z$ axis the top of the valence band is just the state depicted in Fig. 3(a) which is folded back to $\Gamma$ and nonbonding with the tungsten functions]. This absence of interaction accounts for the two bands, one just above and one just below the gap, and for the gap being of similar size as in the sc structure. These two bands do not exhibit dispersion along $\Gamma$-$Z$, since they do not have any lobes pointing away from the $x$-$y$ plane.

Constrained to a smaller volume, the lattice responds by only reducing its dimension in the $z$ direction. It relaxes at its weakest spot: the long W-O bonds along the $z$ direction get compressed [see Fig. 7(a)]. In the electronic structure this translates into the group with the $5d_{xz}$- and $5d_{yz}$-derived bands strongly broadening, and simultaneously approaching the valence-band top (an evident signature of the BAB splitting). However, the $5d_{xy}$-derived band is not affected: it exhibits no dependence on $z$ and along $z$ only does the lattice relax. Therefore, the size of the gap is (nearly) not affected by compression/expansion for this tetragonal structure [Fig. 8(a)].

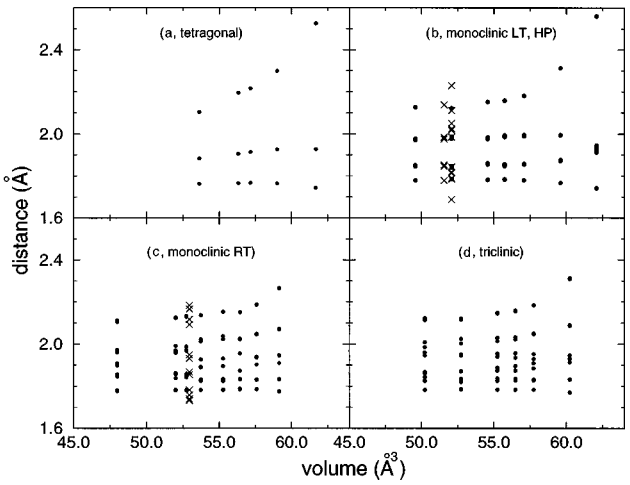


FIG. 7. GGA W-O distances for several crystal phases. The shortest and longest distances are those that are along $z$. Experimental interatomic distances (denoted by $\times$) are shown for the monoclinic RT (Ref. 25), monoclinic LT (Ref. 37), monoclinic HP (Ref. 39).

The distortion from the simple cubic to the tetragonal structure was studied in detail in Ref. 8 with Hartree-Fock (and *a posteriori* GGA corrections for correlation). This study also showed that a strong W-O bond is formed, with increased covalent character, and that it remains largely unaffected under compression. The most significant difference with our study is in the total energies: In Ref. 8 the tetragonal structure was 0.8 eV (per formula unit) more stable than the sc structure. In our study this is only $\sim 0.1$ eV.

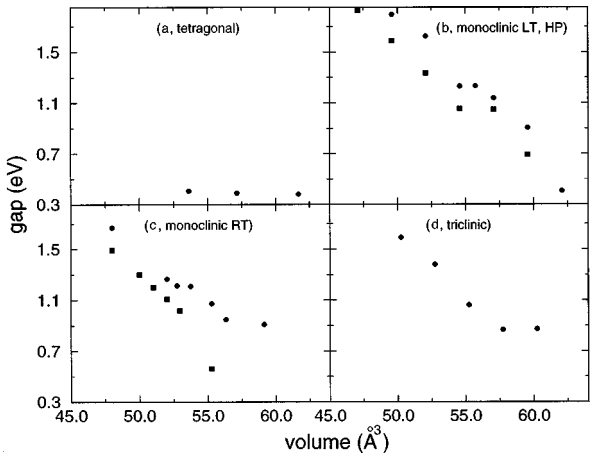


FIG. 8. Gaps as a function of volume. LDA and GGA results are represented by squares and dots, respectively.

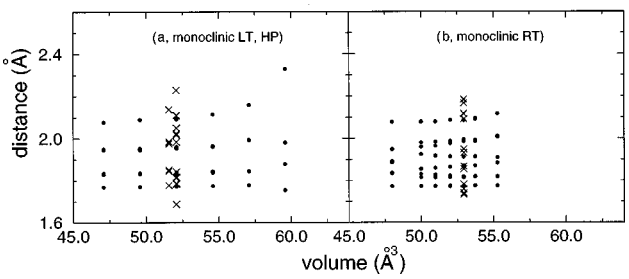


FIG. 9. LDA W-O distances for several crystal phases. See Fig. 7.

### C. Monoclinic low-temperature and high-pressure modifications

Recently experimental diffraction studies have been reported for both the low-temperature monoclinic (Refs. 37 and 38) and high-pressure monoclinic modification.[39] The low-temperature structure is ferroelectrically distorted with space group $Pc$.[29,37,38] The high-pressure structure has space group $P2_1/c$ and was found to occur in the pressure range from 1.2 kbar to 47 kbar (the highest pressure measured). Its derived bulk modulus (at ambient pressure) is 44.5 GPa.[39] In our calculations, starting from either the low-temperature experimental crystal structure or from the high-pressure crystal structure as determined at 5.7 kbar, the minimizations evolved towards very similar structures. The ferroelectric distortion disappears as inversion symmetry develops and the optimized structure has space group $P2_1/c$. This occurs both in the GGA and the LDA. The optimized structure is similar to the experimental high pressure structure [see the bond lengths in Figs. 7(b) and 9(a)]. We can only speculate why the ferroelectric distortion is not stable: e.g., it might be that the distortion is stabilized by a small amount of oxygen deficiency.[37] In the following we discuss the behavior of our computer-generated sample.

The nearest-neighbor (W-O) distances as a function of a volume are shown in Figs. 7(b) and 9(a). Upon expansion, the crystal moves towards the tetragonal structure [compare to Fig. 7(a)]. However, at the equilibrium volume and under compression a long-short splitting occurs for all three directions (the W atom moves off center in the octahedron in all three directions). It is most pronounced along the $z$ direction, whereas in the other two directions it is a bit smaller but of equal magnitude. The splitting is more pronounced than in the monoclinic RT and triclinic modifications (see Secs. III D and III E). In this, the crystal resembles the experimental high-pressure phase, where such a relatively more pronounced splitting was also noticed by the authors of Ref. 39. Indeed, the calculated interatomic distances come rather close to those of the high-pressure phase, but one has to note that of the three short W-O bonds around a W atom, one is very short and two are slightly longer in the calculated structure whereas this is reversed in the experiment.

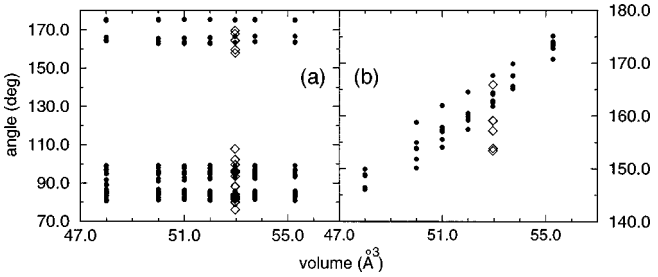


FIG. 10. O-W-O (a) and W-O-W (b) bond angles in the HP monoclinic phase of $WO_3$ for LDA (dots) and experiment (diamonds).

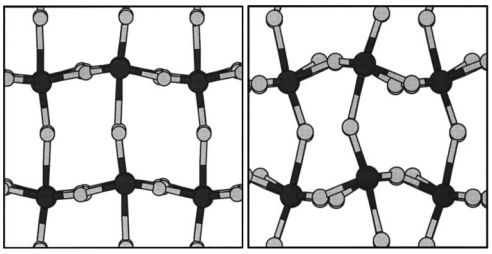

FIG. 11. Illustration of tilting of octahedra in the HP structure. Left: $V=57.1$ Å$^3$, right $V=49.6$ Å$^3$ (GGA structures). The dark and the light spheres represent the W and the O atoms, respectively. A projection onto the $x$-$z$ plane is shown (see Fig. 4), and bonds are drawn between the O and the W atoms. The $z$ axis runs from top to bottom.

The most striking feature of Figs. 7(b) and 9(a) is that all nearest-neighbor distances remain almost constant on compression. Indeed, also the O-W-O angles are hardly affected by the volume change (Fig. 10) so that the shape and size of the octahedra are essentially volume independent. In fact, the crystal responds to pressure only by a rotation of the octahedra relative to one another (Fig. 11, this is a mechanism encountered more often in perovskitelike materials). This is demonstrated by the strong volume dependence of the W-O-W angles, that determine the relative orientation of neighboring octahedra. In fact the stiffness of the W-O-W angles determines how easy it is to compress the crystal. Evidently, this renders compression more easy than expansion (where, ultimately, bonds are being put under strain), as can be seen from the energy versus volume curves in Fig. 12(a). For close-packed materials, there usually is a very steep increase in total energy when the crystal is put under external pressure, but here the compression side of the plot is even flatter than the expansion side. This is most clear for the GGA results, since there the equilibrium volume corresponds best with the crossover of the regime of rotation of octahedra to the regime where bonds are being strained on expansion.

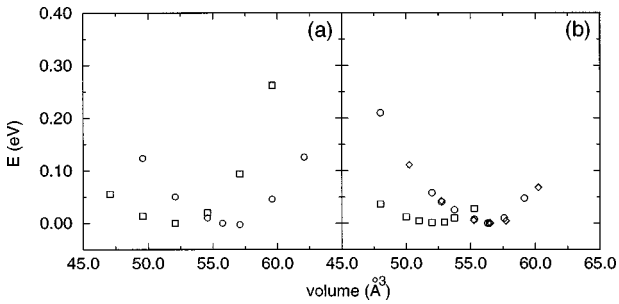


FIG. 12. Total energies as a function of the volume for several crystal structures. Squares and circles depict the LDA and GGA results, respectively. (a) Monoclinic HP/LT structure, and (b) monoclinic RT structure and triclinic structure (GGA only, diamonds).

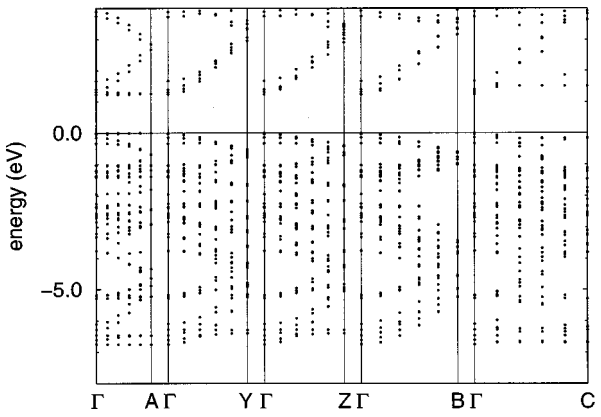


FIG. 13. GGA band-structure fragments of monoclinic LT/HP $WO_3$ for $V=55.8$ Å$^3$ (one formula unit).

The flatness of the energy versus volume curve renders a reliable determination of the bulk modulus difficult. The other complication, most severe for the GGA calculations, is the crossover from one relaxation regime to the other, which renders a Murnaghan-type fit troublesome because $B$ becomes strongly volume dependent. Since there is little point in converging the total energy to within the meV/$WO_3$ level for a closely spaced range of volumes, we can only give approximate numbers for $B$. It is of the same order of magnitude as the low experimental value of 44.5 GPa.[39] We obtained the most reliable number for the monoclinic RT phase (which behaves very similar, see Table II), for which it is accurate to within approximately 10 GPa. The consequences for the electronic structure, notably the semiconductor gap, are large as will be discussed in the following.

The electronic gaps have not been calculated from a full BZ scan, but only a few **k** points have been considered. Apart from the points used in the self-consistent calculation (see Sec. II) these were those in the centers of the zone-boundary faces ($A,B,Z,Y$) and $\Gamma$ (see also Sec. III D). The top of the valence band and the minimum of the conduction band were always found at $\Gamma$ and/or $A$ (GGA and LDA), consistent with the tetragonal and sc structures. In fact, close to the gap, the states at $A$ are almost degenerate with those at $\Gamma$. As an extra check, we carried out a less restricted BZ scan for the GGA ''equilibrium'' volume. The scan involved several points on the lines going out from $\Gamma$ to the $A$, $B$, $Z$, and $Y$ points and in addition, the point $C$ (which is in the middle of an edge of the BZ). The band structure along these lines is shown in Fig. 13. The states near the gap have no dispersion along $\Gamma$-$A$ and indeed the gap is located on this line.

The gap has increased appreciably compared to the tetragonal structure [Fig. 8(b)]. This was to be expected since a BAB splitting between W $5d_{xy}$ orbitals and O $2p_x$ and $2p_y$ has now also become possible. Also evident from Fig. 8(b) is a sizeable increase of the gap with compression. This can be rationalized from a comparison with the behavior of the sc phase. There the different wave-function characters of the valence and conduction bands caused these bands to move apart under compression. This resulted in a widening of the gap in principle. However, this tendency was counteracted by the band broadening, and the combined result was a weak increase of the gap with compression. For the present monoclinic structure, the octahedra remain rigid entities: The intraoctahedral O-O distances do not change. This should frustrate the band broadening and, therefore, the gap increases more dramatically with compression. Of course, this picture is somewhat oversimplified. Even in the sc structure, tungsten and oxygen states mix and the bandwidth is, in part, also determined by W-O interactions. However, the states next to the gap are of pure tungsten or oxygen character so that this additional interaction does not affect the behavior near the gap. For the monoclinic structure, however, these states do mix, causing the BAB splitting that further widens the gap. Since the octahedra remain rigid, the BAB splitting is not much affected by the compression and should not much affect the mechanism causing the increase of the gap.

### D. Monoclinic room-temperature modification

The unit cell of this structure is twice the size of the unit cell of the monoclinic low-temperature structure ($8\times WO_3$): it contains eight corner-sharing oxygen octahedra in a nearly cubic arrangement.

Again, the equilibrium volume is overestimated by the GGA (see Table II). The LDA removes most of this discrepancy. The shortest interatomic distances are somewhat overestimated and the longest a bit underestimated [see Figs. 7(c) and 9(b)]. GGA distances are slightly longer than the LDA distances.

We have calculated the gap as a function of volume. Because of the considerable computational cost involved, we have only calculated the eigenvalues at a few **k** points. These are those that were used in the structure optimization, some **k** points at the BZ boundary ($A,B,Y,Z$) and $\Gamma$. With such a limited selection of points, it is possible (in principle) to overlook the extrema of the valence and the conduction bands, and thus to overestimate the gap. To test the validity of the usage of this limited selection, we have carried out a very thorough BZ scan with the localized-spherical-wave (LSW) method (Ref. 30) and compared the results with those of the pseudopotential calculation. For this, experimental atomic positions were used.[25] The LSW method is much faster than any plane-wave method and therefore allows for a very thorough scan of the BZ. Since it makes a shape approximation to the potential (atomic sphere approximation), care needs to be taken in setting up correctly the atomic spheres. Their radii were tuned such as to reproduce as well as possible the band structure of the sc-$WO_3$, which was calculated with the pseudopotential method. However, since the monoclinic structure is quite distorted, some additional optimization of the atomic sphere radii was found necessary in order to obtain a good filling of space. The LSW calculation was performed in the LDA only, since differences between LDA and GGA band structures are small in general. It is found that the extremes of the valence and conduction bands indeed are located at the special points that we consider in the pseudopotential calculation, giving confidence that we obtain accurate numbers for the gap. Moreover, both with the LSW and the pseudopotential calculation, the gap is direct and occurs at $\Gamma$. With LSW it is 1.58 eV, to be compared with the 1.68 eV obtained with the pseudopotential calculation (both calculations using the experimental positions).

At the calculated GGA and LDA equilibrium volumes the size of the GGA and the LDA gap is 0.90 eV and 1.1 eV,

respectively. It is direct and observed at $\Gamma$. Moreover, the bottom of the conduction band is degenerate between $\Gamma$ and *B*. The top of the valence band seems rather flat, and degeneracies with $\Gamma$ occur at different parts of the BZ for GGA and LDA. This difference can be attributed to the volume difference (see below). Koffyberg *et al.* have measured the gap and obtained a value of 2.62 eV (Ref. 31), for an indirect transition. Underestimation of the gap is a frequently occurring feature of LDA/GGA methods.[32]

We find a direct gap where Koffyberg *et al.* have obtained an indirect gap. In the following we argue that this distinction could be artificial, i.e., under suitable circumstances a direct gap can give rise to the same experimental properties that led Koffyberg *et al.* to conclude that they observed an indirect gap.[31] In their experiment, the character of the gap is indirectly determined from the kind of energy scaling of the optical absorption.[33] The optical absorption is proportional to a BZ integration of matrix elements of the momentum operator (see, e.g., Ref. 34), and whether these matrix elements belong to direct or indirect (and/or forbidden) transitions determines the power of the energy scaling. Thus one has a means to extract information on the nature of the gap from measurements of the optical absorption as function of energy. However, there is some arbitrariness in this procedure. For example, by taking a larger unit cell, matrix elements formerly corresponding to indirect transitions can, by means of back folding, transform into direct transitions, without any of the optical properties being affected. An example is the sc structure of $WO_3$: we saw already that if a nonprimitive unit cell is taken (of the same dimensions as the tetragonal cell) the gap folds back to $\Gamma$ and becomes direct (without any matrix element being affected). Of course, this example is a bit artificial, but one can imagine something similar happening also if a distortion of the lattice necessitates an enlargement of the unit cell. In that case not only do several states fold onto the same spot in the new BZ, but also levels shift and matrix elements change. However, the extent to which states are affected may be very different. Some may hardly change and just fold back whereas others (involved in the interactions driving the distortion) may change significantly in some respects. The result for the scaling of the matrix elements is not *a priori* clear, and it is well possible that some matrix elements of direct, but formerly indirect, transitions have not appreciably changed. So they should appear with the scaling ''appropriate'' for indirect transitions in an optical experiment. A similar back folding procedure can be envisaged for the present structure: the monoclinic BZ is very close to a cubic supercell into which eight simple cubic cells can be fitted. The *M*, *X*, and *R* point of these cells just fold into the $\Gamma$ point of the cubic supercell (but folding itself does not alter the optical properties). The deformations in the larger monoclinic unit cell can modify the transition probabilities but not necessarily to the extent that the formerly indirect transitions develop into a direct (allowed) absorption channel of appreciable intensity.

For the LDA sample, upon compression the gap remains between $\Gamma$ and *Y* in the valence band and $\Gamma$ and *B* in the conduction band ($\Gamma$-$Y \rightarrow \Gamma$-$B$), i.e., the transition remains direct. Only at very high compression ($V = 48$ Å$^3$) the transition becomes indirect ($Y \rightarrow \Gamma$-$B$). Upon expansion the gap becomes located on $\Gamma$-*B* at approximately the GGA equilibrium volume. The GGA sample behaves similarly. Its equilibrium volume is larger, so for it the gap is along $\Gamma$-*B*.

Figures 7(c) and 9(b) show the W-O distances for the monoclinic RT structure. A splitting into long and short bonds is again evident, but now it is of different magnitude for the *x*, *y*, and *z* directions. In fact, for several O-W-O triplets it is weaker than the splitting found in the monoclinic LT structure [Figs. 7(b) and 9(a)]. Therefore the BAB splitting is less effective for several W-O pairs, which helps to understand why the gap (at the equilibrium volume) in the RT structure is slightly smaller than in the LT structure.

Starting from the equilibrium structure and expanding, a few W-O bonds are strongly elongated (along *z* and one of the bonds in the *x*,*y* plane). This should lead to an increase of the distance between several bands in the conduction-band and valence-band regions, because of the BAB mechanism. However, this mechanism is very weak for at least one direction, which may account for the weak volume dependence of the gap that we find upon expansion [Fig. 8(a), GGA].

When the crystal is compressed, something very different happens: The longest bonds remain just as long, the shortest bonds do not shorten. A mechanism similar to what happens for the compressed monoclinic LT phase takes place: The rigid octahedra start to tilt relative to one another. This correlates with a considerable widening of the gap (Fig. 8). Just as for the smaller monoclinic cell, we interpret this as a frustration of band broadening.

Comparing the LDA and the GGA calculations, we see again that the GGA overestimates the equilibrium volume and that W-O distances are a bit larger within the GGA. Both LDA and GGA do not entirely reproduce the large range of W-O distances found in the experiment. In fact, the shortest calculated distances are too long and the longest are too short. Moreover, the size of the (rigid) octahedra is almost identical for LDA and GGA and very close to the experimental size. What really results in the difference in equilibrium volumes is the difference in the interactions between the octahedra, i.e., the different ''angular potentials.'' Evidently for this type of interaction the LDA gives better results.

Another advantage of the LDA is the stability of the local minima. In Fig. 9(b) the variation of interatomic distances is very smooth as a function of volume, whereas for the GGA various jumps occur. In fact, the GGA results ''oscillate'' between two nearly degenerate structures. We found this in the LDA calculations also, but there we could separate out both structures. The ''spurious'' structure is probably a precursor of the triclinic phase that cannot fully develop in the monoclinic cell. Its total energy differs by approximately 1 meV per formula unit. Such a small energy difference, being beyond the accuracy of our calculation, is not meaningful anymore.

### E. Triclinic modification

The triclinic structure (Ref. 35) differs only slightly from the monoclinic RT one. It also has a nearly cubic supercell consisting of eight corner-sharing $WO_3$ octahedra. The unit-cell dimensions are very similar to those of the monoclinic RT structure (Table III). Splitting into long and short bonds again occurs for the *x*, *y*, and *z* directions [Fig. 7(d)] and again to a different extent along all three independent directions. The octahedra tilt relative to one another resulting in

zigzag chains of tungsten oxygen bonds forming in all directions. Again, this also holds for the monoclinic RT structure. The essential difference is in how these zigzag chains are ordered relative to one another. The chains always have four nearest-neighbor chains. In the triclinic structure, these always zigzag in the direction opposite to the zigzag of the central chain: they are ''180° out of phase.'' In the monoclinic structure this does not hold for one of the directions. For that particular direction only two of the neighboring chains are out of phase, the others are not. This difference is very subtle, and it needs to be stressed that chains that are in phase are not identical to the central chain, they merely have the same ''phase'' in the different zigzag distortions.

## IV. DISCUSSION AND CONCLUSIONS

In this paper we have studied several crystal structures of tungsten trioxide that are distortions of the artificial perovskitelike simple cubic structure of Fig. 1. Whereas the electronic structure of this artificial sc-$WO_3$ can be understood in terms of a simple, essentially ionic picture, for the other structures covalent effects are of much more significance. The alternation of long and short W-O bonds reflects the occurrence of a bonding–anti-bonding splitting that helps widen the gap by more than a factor of 2. Such a mechanism can only be frustrated if breakup into long and short bonds remains limited to just one direction. This occurs for the tetragonal modification, where the $5d_{xy}$ tungsten states cannot develop a substantial overlap with neither the $p_x$ nor the $p_y$ ''orbitals'' of their oxygen neighbors. Consistently, only for this structure the gap is of similar size as for the sc structure. In the other structures such a long-short splitting occurs in all directions, although not to the same degree in each of them.

The long-short splitting can be decomposed into an elongation of the oxygen octahedra and a movement of the tungsten atoms away from the centers of the octahedra. In fact, the oxygen octahedra are part of an ''orthorhombic lattice'' that is not much affected by the occurrence of the long-short splitting. This can change if the lattice is put under pressure. In the more complicated structures the octahedra can tilt relative to one another. As a result, no compression is put on the W-O bonds and the octahedra tilt as nearly rigid entities. This results in a very flat energy versus volume curve and thus a low bulk modulus. This is corroborated by the diamond-anvil-cell experiments of Ref. 39.

The structural relaxation mechanisms correlate strongly with the electronic structure. In the sc structure the different degree of localization of the (oxygen-derived) valence-band and (tungsten-derived) conduction-band states results in a tendency of the valence and conduction bands to move apart under compression. However, band broadening reduces this effect significantly, leading to a merely weak volume dependence of the gap.

In the monoclinic and triclinic structures this band broadening is much weaker (at least for the valence band) and, in consequence, a strong increase of the gap with compression is observed. The reason for this absence of broadening lies in the structure relaxation: nearest- (W-O) and next-nearest- (O-O) neighbor distances do not change on average, thus preventing the broadening. [In addition to this comes the interaction of BAB-type, absent in the sc structure, that results in an increase of the gap at equilibrium (ambient) conditions.] We expect this mechanism to break down if the shortest interoctahedral O-O distances becomes equal to the intra-octahedral O-O distances. This happens in our most compressed structures (where the gap becomes indirect also). Further compression should result in a different behavior.

The GGA gives somewhat larger W-O bonds than the LDA, which is in agreement with the general trend. However, both result in octahedra of similar size. The equilibrium volume is determined by the (bending) interactions of the W-O-W angles (balancing the electrostatics, of course) and these are evidently better described by the LDA.

## ACKNOWLEDGMENTS

We acknowledge stimulating discussions with Dr. E.P. Boonekamp, Dr. T.J. Vink, and Professor L.F. Feiner. This work was part of the research program of the Stichting for Fundamenteel Onderzoek der Materie (FOM) with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

[1] C. G. Granqvist, *Handbook of Inorganic Electrochromic Materials* (Elsevier, Amsterdam, 1995).
[2] L. Kopp, B. N. Harmon, and S. H. Liu, Solid State Commun. **22**, 677 (1977).
[3] D. W. Bullett, J. Phys. C **16**, 2197 (1983).
[4] D. W. Bullett, Solid State Commun. **46**, 575 (1983).
[5] C.-G. Zhan and F. Zheng, J. Mol. Struct.: THEOCHEM **285**, 89 (1993).
[6] A. Stashans and S. Lunell, Int. J. Quantum Chem. **63**, 729 (1997).
[7] N. E. Christensen and A. R. Mackintosh, Phys. Rev. B **35**, 8246 (1987).
[8] F. Corà, A. Patel, N. M. Harrison, R. Dovesi, and C. R. A. Catlow, J. Am. Chem. Soc. **118**, 12 174 (1996).
[9] M. G. Stachiotti, F. Corà, C. R. A. Catlow, and C. O. Rodriguez, Phys. Rev. B **55**, 7508 (1997); F. Corà, M. G. Stachiotti, C. R. A. Catlow, and C. O. Rodriguez, J. Phys. Chem. B **101**, 3945 (1997).
[10] A. Hjelm, C. G. Granqvist, and J. M. Wills, Phys. Rev. B **54**, 2436 (1996).
[11] *Gmelin Handbuch der Anorganische Chemie, Wolfram Ergänzungsband B2, Oxide 1979* (Springer, Berlin, 1979).
[12] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993); **49**, 14 251 (1994).
[13] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
[14] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11 169 (1996).
[15] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).
[16] G. Kresse and J. Hafner, J. Phys.: Condens. Matter **6**, 8245 (1994).
[17] S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).

[18] R. D. King-Smith, M. C. Payne, and J. S. Lin, Phys. Rev. B **44**, 13 063 (1991).
[19] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[20] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
[21] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).
[22] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 1588 (1976).
[23] P. E. Blöchl, O. Jepsen, and O. K. Andersen, Phys. Rev. B **49**, 16 223 (1994).
[24] B. O. Loopstra and P. Boldrini, Acta Crystallogr. **21**, 158 (1966).
[25] B. O. Loopstra and H. M. Rietveld, Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. **B25**, 1420 (1969).
[26] A. K. McMahan and R. C. Albers, Phys. Rev. Lett. **49**, 1198 (1982).
[27] W. L. Kehl, R. G. Hay, and D. Wahl, J. Appl. Phys. **23**, 212 (1952).
[28] E. Salje and K. Viswanathan, Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. **A31**, 356 (1975).
[29] E. Salje, Ferroelectrics **12**, 215 (1976).
[30] H. van Leuken, A. Lodder, M. T. Czyżyk, F. Springelkamp, and R. A. de Groot, Phys. Rev. B **41**, 5613 (1990).
[31] F. P. Koffyberg, K. Dwight, and A. Wold, Solid State Commun. **30**, 422 (1979).
[32] Note that the underestimation of the gap originates from two effects: First, the GGA underestimates the gap with the atoms at their experimental positions (for this gap we find 1.62 eV). Second, we calculate the gap in for the optimized GGA volume, giving rise to an additional decrease (from 1.62 to 0.90 eV).
[33] M. A. Butler, J. Appl. Phys. **48**, 1914 (1977).
[34] J. N. Hodgson, *Optical Absorption and Dispersion in Solids* (Butler and Tanner, London, 1970).
[35] S. Tanisaki, J. Phys. Soc. Jpn. **15**, 566 (1960).
[36] E. A. Nikitina and A. S. Kokurina, Zh. Obshch. Khim. **21**, 1940 (1951).
[37] E. K. H. Salje, S. Rehmann, F. Pobell, D. Morris, K. S. Knight, T. Herrmannsdörfer, and M. T. Dove, J. Phys.: Condens. Matter **9**, 6564 (1997).
[38] P. M. Woodward, A. W. Sleight, and T. Vogt, J. Solid State Chem. **131**, 9 (1997).
[39] Y. Xu, S. Carlson, and R. Norrestam, J. Solid State Chem. **132**, 123 (1997).