npj | Computational Materials

www.nature.com/npjcompumats

ARTICLE OPEN

Check for updates

Wannier–Koopmans method calculations for transition metal oxide band gaps

Mouyi Weng[1], Feng Pan[1]✉ and Lin-Wang Wang[2]✉

The widely used density functional theory (DFT) has a major drawback of underestimating the band gaps of materials. Wannier–Koopmans method (WKM) was recently developed for band gap calculations with accuracy on a par with more complicated methods. WKM has been tested for main group covalent semiconductors, alkali halides, 2D materials, and organic crystals. Here we apply the WKM to another interesting type of material system: the transition metal (TM) oxides. TM oxides can be classified as either with d^0 or d^{10} closed shell occupancy or partially occupied open shell configuration, and the latter is known to be strongly correlated Mott insulators. We found that, while WKM provides adequate band gaps for the d^0 and d^{10} TM oxides, it fails to provide correct band gaps for the group with partially occupied d states. This issue is also found in other mean-field approaches like the GW calculations. We believe that the problem comes from a strong interaction between the occupied and unoccupied d-state Wannier functions in a partially occupied d-state system. We also found that, for pseudopotential calculations including deep core levels, it is necessary to remove the electron densities of these deep core levels in the Hartree and exchange–correlation energy functional when calculating the WKM correction parameters for the d-state Wannier functions.

npj Computational Materials (2020)6:33 ; https://doi.org/10.1038/s41524-020-0302-0

INTRODUCTION

Transition metal (TM) oxide is one of the most important class of materials used in electronics, solar cells, catalysts, batteries, and many other devices. Band gap of the material is an essential characteristic for these applications. It is, thus, of paramount importance to be able to predict the band gaps of the material by using first principle calculations. However, it is well known that the density functional theory (DFT) approximated by local density approximation (LDA)[1] or generalized gradient approximation (GGA)[2] and implemented with Kohn–Sham equations significantly underestimates their band gaps due to the lack of discontinuity in their functional derivatives. There are methods utilizing empirical parameters to derive correct band gaps, such as meta-GGA[3], HSE06[4], B3LYP[5,6], and Hubbard U methods[7]. However, the empirical nature of these methods leaves a lot to be desired. A more standard first principle approach requiring no empirical parameters to calculate the band gap is the many-body perturbation GW method. However, even for the GW method[8], getting the correct TM oxide band gap is a challenge. Not only it can have different flavors in regard to the self-consistency, including G_0W_0, GW_0, and GW (here 0 means no self-consistency), which can give different results, but also the initial input wave function and eigenenergies can also influence the final results for the not fully self-consistent calculations. Moreover, there are also challenges for numerical implementations (e.g., plane wave basis versus localized basis set) and other approximations (e.g., plasmon pole approximation for the dielectric function). Besides, some of the calculations require a large number of conduction bands to yield fully converged results, which can lead to debates about the accuracy of the GW method. Thus, even if fully converged GW method might be good, there could be many problems in practical use of its calculations either due to numerical implementation issues or convergence issues. For example, for ZnO, the experimentally measured band gap is 3.43 eV[9], while the reported GW calculated band gap ranges from 2.1 to 3.9 eV[10]. The situation is even worse for TM oxides with partially occupied d states. For rock salt MnO system, the experimentally measured band gap is believed to be 4.0 eV[11]. Falee et al. reported an all-electron self-consistent GW result of 3.5 eV in 2004[12]. Scheffler et al. reported a G_0W_0@LDA + U result of 2.34 eV and a GW_0@LDA + U result of 2.57 eV in 2010[13]. Lany reported a GW in the random-phase approximation band gap of 3.81 eV and GW with local-field effects and empirical V_d potential band gap of 3.36 eV in 2013[14]. Manousakis et al. reported converged self-consistent GW result of 4.39 eV in 2015[15]. Jiang reported a G_0W_0 result with high-energy local orbitals and linearized augmented plane waves of 3.32 eV, and the GW_0 result is 3.69 eV in 2018[16]. Moreover, Carter tested G_0W_0, GW_0, and GW, on PBE, LDA + U, PBE + U, HSE06, and PBE0 for Fe_2O_3 in 2011, and the results vary from 1.3 to 4.8 eV compared with an experimentally measured band gap of around 2.6 eV[17]. On top of all these, GW calculation can be expensive, especially when a large number of conduction band states are needed or when the system size is large. Given this situation, it will be very interesting to find whether there are alternative parameter-free methods to predict the TM oxide band gaps.

There have been many prior works in using Koopmans condition to improve the accuracy of band gap of system especially for atoms and small molecules. For example, Dabo et al.[18] proposed a generalized Koopmans theory to remove not only the electrons from the highest valence states but also the electrons from other states to correct the energies of these states. They applied this method to atoms and small molecules and achieved excellent agreement with the experimental results. Cococcioni et al.[19] used Koopmans condition to fix the U parameter in DFT + U calculations. Kraisler et al.[20] showed that it is possible to use Koopmans condition to correct any exchange-correlation functional, restore the discontinuity of its derivative, and used that to predict the electron affinities and ionization

[1]School of Advanced Materials, Peking University Shenzhen Graduate School, 518055 Shenzhen, People's Republic of China. [2]Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. ✉email: panfeng@pkusz.edu.cn; lwwang@lbl.gov

Published in partnership with the Shanghai Institute of Ceramics of the Chinese Academy of Sciences

npj nature partner journals

energies of atoms and small molecules. Over the past few years, we have developed a Wannier–Koopmans Method (WKM) which can yield accurate band gaps on main group covalent semiconductors[21], alkali halides[22], 2D materials[23], and organic crystals[24]. For these tested materials, the accuracy of WKM is similar to that of GW method. In our approach, Wannier functions and the Koopmans condition are combined to calculate the band gap of bulk systems. Similar methods have also been adopted in recent years. For example, Yang et al. proposed a localized orbital scaling correction framework[25]. Instead of implementing a post-DFT correction method like ours, they applied it during the self-consistent (SCF) calculations. However, the main purpose is similar: to use a localized function in compliance with the Koopmans condition to correct the DFT band gap. Marzari et al. reported a Koopmans-compliant spectral functional method[26]. For the conduction band, they used an approach similar to ours with localized Wannier functions. For valence band, they used the self-interaction correction (SIC) function to derive the localized orbital. Since SIC localized orbital and maximally localized Wannier functions are very similar, their results are also very similar to WKM results. Given all the interest in this approach, it is thus important to test the limit of WKM. One class of important materials which have not yet been systematically tested is the TM oxides. The TM oxides can be classified into two groups. One group is with d^0 or d^{10} configurations for the TM elements, and the other group is with partially occupied d states. The second group of TM oxides have long been regarded as strongly correlated Mott insulators, which are difficult to be described using mean field-like approaches (e.g., DFT, SIC, or even GW). It will, thus, be interesting to see whether there are qualitative differences between these two groups of TM oxides when the WKM method is applied.

In our following calculations, we found that the WKM method works well for the d^0 or d^{10} close shell systems, whereas it performs badly for the partially occupied open shell configurations. The main cause of this problem for the open shell system is due to a strong interaction between the occupied and unoccupied d-Wannier functions as we will show later. This is perhaps compounded by the fact that the DFT ground-state properties (e.g., the magnetic moment) of the open shell systems can also be significantly wrong.

WKM is an extension of the ΔDFT method[27]. In the traditional ΔDFT method, the band gap is calculated by using the difference between the electron affinity (EA) energy and the ionization energy (IE). EA and IE can be calculated by using self-consistent ground-state energy $E(N+1)$, $E(N)$, and $E(N-1)$. Here, N is the number of electrons in the neutral system, and $N+1$ and $N-1$ indicate that the system has one more or one less electron, respectively. EA can be expressed as $E(N+1)-E(N)$ and IE as $E(N)-E(N-1)$. Although this method works well for isolated molecules, the density change caused by adding or subtracting one electron is infinitesimally small in an extended system. As a result, according to Janak's theory[28], the total energy difference is the same as Kohn–Sham orbital eigenenergy. To overcome this problem, we added an electron into a localized Wannier function instead of the extended Kohn–Sham orbitals. The resulting LDA total energy with partial occupation s_w of the Wannier function ϕ_w can be denoted as $E_{\mathrm{LDA}}(\{s_w\})$. This function is not a linear function of s_w between 0 and 1. However, in many-body quantum mechanics, the total energy of a system with a fractional number of electrons can be defined as a statistical mixture of the N electron and $N\pm1$ electron state. This leads to a linear segment total energy function of s_w. This linear segment property is also called Koopmans condition. More deeply, this means that the system (or a relatively isolated part of the system) likes to occupy an integer number of electrons, instead of partial number of electrons. This would be a consequence of a grand canonical minimization using the linear segment total energy function with a given external Fermi energy (a consequence of a linear programming optimization). The integer number of electrons in a given system is also a property of many-body electron wave treatment of the system. Finally, to satisfy the Koopmans condition, one compensation term $E_w(s_w)$ can be added to the total energy expression. As a result, the WKM total energy can be expressed as

$$E_{\mathrm{WKM}}(\{s_w\}) = E_{\mathrm{LDA}}(\{s_w\}) + \sum_w E_w(s_w)\,. \tag{1}$$

Here w indicates a set of Wannier functions that are orthogonal to each other. s_w $(0 < s_w < 1)$ indicates the occupation number of this Wannier function ϕ_w. $E_w(s_w)$ is a simple function of s_w obtained by requiring that $E_{\mathrm{WKM}}(\{s_w\})$ is a linear function of s_w. During the LDA calculation, we add or remove the electron from ϕ_w of one spin channel and all the other orbitals in this spin channel should be orthogonal to this Wannier function ϕ_w. All the other orbitals (except this one Wannier function) are variationally changed to minimize the total energy, which results in the ground-state energy $E_{\mathrm{LDA}}(\{s_w\})$.

A simple analytical expression of $E_w(s_w)$ can be written as

$$E_w(s_w) = \lambda_w s_w(1 - s_w) \tag{2}$$

The λ_w can be determined from $E_{\mathrm{LDA}}(\{s_w\})$ (to make $E_{\mathrm{WKM}}(\{s_w\})$ a straight line versus s_w). The $E_{\mathrm{LDA}}(\{s_w\})$ calculation is done using a supercell that should be large enough to remove the image interactions between neighboring supercells. In our calculations, we used sufficiently large supercells to reduce the image charge problem. Based on a test in our previous work[23], the image charge effect on band gap is likely to be in the order of 0.01 eV when the distance between Wannier functions from the neighboring supercell is >10 Å. Note, in the future, image interaction correction techniques can be used to reduce the needed supercell sizes, like that in a charged impurity defect calculation.

After doing the self-consistent calculation for $E_{\mathrm{LDA}}(\{s_w\})$ with a Wannier function ϕ_w, we can obtain λ_w. For a Kohn–Sham orbital ψ_i, we can have the modified eigenenergy ε_i with $[H_{\mathrm{LDA}} + \sum_w \lambda_w |\phi_w\rangle\langle\phi_w|]\psi_i = \varepsilon_i\psi_i$. More details of this method can be found in ref. [21].

RESULTS AND DISCUSSION

WKM calculations for d^0 and d^{10} transition metal oxide

Three main types of compounds are considered for the d^0 and d^{10} systems in this work, which are spinel compounds, perovskites compounds, and several binary TM oxide compounds. The four spinel structures are $ZnGa_2O_4$, $CdGa_2O_4$, $CdIn_2O_4$, and $MgIn_2O_4$ (structural details are shown in Supplementary Fig. 1 and Supplementary Table 1). The six perovskites are $CaTiO_3$, $CaZrO_3$, $CaHfO_3$, $SrTiO_3$, $SrZrO_3$, and $SrHfO_3$ (structural details are shown in Supplementary Fig. 2 and Supplementary Table 2). The six transition binary TM oxides are Sc_2O_3, anatase TiO_2, rutile TiO_2, ZrO_2, Cu_2O, and ZnO (structural details are shown in Supplementary Fig. 3 and Supplementary Note 1). We used lattice constants from the Materials Project relaxed result for all these systems[29]. The details of these structures are listed in the Supplementary Materials.

WKM calculations were performed in two different sets of norm conserving pseudopotentials: the FHI[30–34] pseudopotential and the recent SG15 pseudopotential[35,36]. The FHI pseudopotential does not include semicore states, while the SG15 includes the semicore states. Their detailed valence electron configurations are included in Supplementary Table 3. In general, the semicores are defined as the subshell levels that are not in the row of the periodic table where this element resides. For example, for the fourth-row elements (e.g., Fe), the 3s and 3p levels will be semicores. However, we have also taken O-2s level as the

semicore as it is often isolated 10 eV below the other part of valence bands near the band gap.

For the LDA calculations, the SG15 and FHI pseudopotential yield similar band gaps. Here, for clarity, we will call the straight forward WKM procedure described in ref. [21], as the "normal" WKM procedure. This is to distinguish it from a modified WKM procedure where the semicore electron charge density is excluded from the Hartree and exchange–correlation energy (E_{hxc}) calculations (as will be discussed later). The normal WKM calculation shows that, in most d^0 and d^{10} cases, the SG15 gives a larger WKM band gap than FHI, and it can be 1 eV too big compared with the experiment. The results are shown in Fig. 1 labeled as the blue squares and SG15-WKM. In stark contrast, the FHI–WKM results are close to the experimental values as shown in red circle.

We believe that the difference between the SG15 and FHI–WKM band gaps comes from the interaction between the semicore-level electrons and the d-orbital Wannier function electrons via the Hartree and exchange–correlation energy. In a sense, if we believe that the band gap correction should come from the valence band properties, then the correction parameter like λ should also be able to be calculated by the valence bands alone. If the inclusion of the semicore level in the calculation makes a major difference, then we should trust the original calculation based on the valence bands alone. This argument leads us to the following procedure to exclude the semicore electron charge density from the E_{hxc} calculation.

The original Hartree and exchange-correlation energy E_{hxc} is calculated as

$$E_{hxc} = E_{hxc}(\rho_c + \rho_v + s_w\rho_w). \quad (3)$$

Here ρ_c, ρ_v, and ρ_w stand for semicore, valence, and Wannier orbital charge densities, respectively. Now, in a revised version, we have:

$$E_{hxc_{fr}} = E_{hxc}(\rho_c + \rho_v) + E_{hxc}(\rho_v + s_w\rho_w) - E_{hxc}(\rho_v). \quad (4)$$

This expression is used when calculating $E_{LDA}(\{s_w\})$ of Eq. (1) as well as for total energy minimization when ϕ_w is partially occupied with s_w in the WKM SG15 pseudopotential calculations. During the SCF iterations, we have also kept the semicore-level wave function ψ_c unchanged from their original LDA wave function inputs. Physically, this procedure is equivalent to excluding the semicore levels from participating in the screening of the Wannier function charge. Usually, one might think such screening should reduce the charge self-interaction, thus reducing the amplitude of λ_w. Surprisingly, in reality, we found that including such semicore can increase the amplitude of λ_w, probably due to a highly nonlinear exchange-correlation energy contribution. This might come from the error of the LDA exchange-correlation function. It thus also partially justifies the removal of the semicore term from the E_{hxc} term. At this stage, however, we do treat this procedure of removing the semicore level from exchange-correlation energy as an empirical procedure.

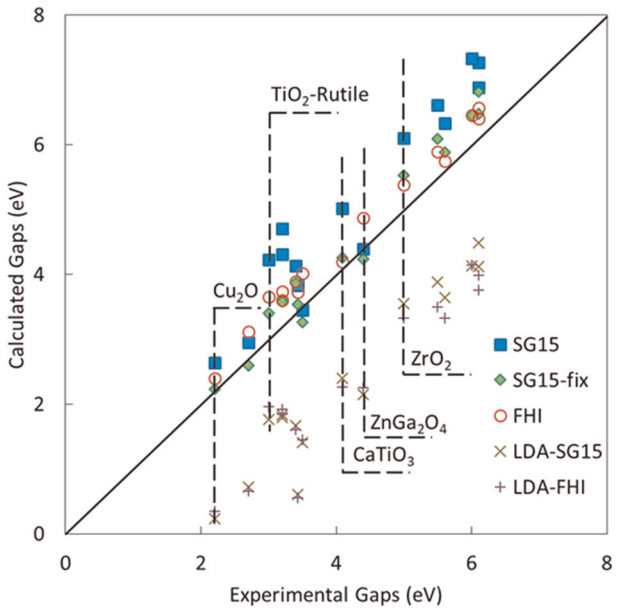

Fig. 1 Band Gap comparison between WKM/LDA calculations and experimental results. Comparing the experimental band gaps against band gap calculated by WKM in SG15 pseudopotential (labeled as SG15-WKM), SG15 pseudopotential with core-level fixation (labeled as SG15-WKM-fix1), FHI pseudopotential (labeled as FHI–WKM) and LDA with SG15 (labeled as SG15-LDA) and FHI (labeled as FHI–LDA) pseudopotential.

Let us now use anatase TiO_2 as an example to discuss this procedure in more detail. The LDA band structure and projected density of states of anatase TiO_2 is shown in Fig. 2a, b. The LDA band gap of anatase TiO_2 is 1.8 eV. The calculations are performed using SG15 pseudopotentials. In the SG15 pseudopotentials for Ti and O, we have several semicore states. Since Ti is a fourth-row element, the Ti-3s and Ti-3p levels are treated as the semicore, and their corresponding band structure and density of states are shown in Fig. 2a, b. We then generate Wannier functions for anatase TiO_2 by wannier90 code[37]. For valance band part, we used initial guess of p orbital on O atoms; for conduction band part, we used initial guess of d orbital of Ti atoms. The Wannier functions with the maximum project weights on conduction band minimum (CBM) and valence band maximum (VBM) are shown in Fig. 3 within the supercell used for the λ_w calculation.

We then remove different numbers of semicore electrons in the E_{hxc} calculation. The results are shown in Table 1. The experimental band gap for anatase TiO_2 is 3.2 eV[38]. The more semicore states are removed in the E_{hxc}, and the smaller band gaps are obtained in WKM. We also notice that the main difference comes from the λ_w from the conduction band.

To test whether this semicore removal from E_{hxc} procedure is necessary for other systems where there is no d-state Wannier functions, we also calculated the AlAs. The experimental band gap for AlAs is 2.2 eV[39]. In the Wannier generation, we used As-p orbital for valance bands and Al-s for conduction bands. The results are shown in Table 2. As we can see, for this system, the effects of the semicore level are rather small. Thus, we conclude that, it is not necessarily to have this E_{hxc} semicore-level removal procedure when calculating the λ_w of non-d-state Wannier functions.

We also used GGA for E_{hxc} in our calculation. A trend similar to LDA is shown in Supplementary Table 4. Lastly, we expect the same semicore removal procedure should also be used for all-electron calculations.

The results of using this E_{hxc} semicore-level removal procedure for the d^0 and d^{10} TM oxides using SG15 pseudopotentials are shown in Fig. 1 as green diamond (labeled as SG15-WKM-fix). Now, we can see that the SG15 and FHI results are very similar, and they are all close to the experimental band gaps. The band gap values of SG15 calculation, SG15-WKM-fix calculation, FHI calculation, and experimental result can be found in Supplementary Tables 5–8, respectively. We, thus, conclude that the WKM works well for the d^0 and d^{10} TMs after this core-level removal procedure in E_{hxc}. It is interesting to note that most of the WKM-predicted band gaps are still slightly larger than the experimental results. It will be interesting to test in the future, whether the electron–phonon coupling effect, which tends to reduce the band gap[40], can bring the theoretical results even closer to the experimental values.

WKM calculations for partial occupied d-orbital TM oxide

Having finished the calculations for d^0 and d^{10} TM oxides, we now turn our attention to the partially occupied d-state oxides. In this

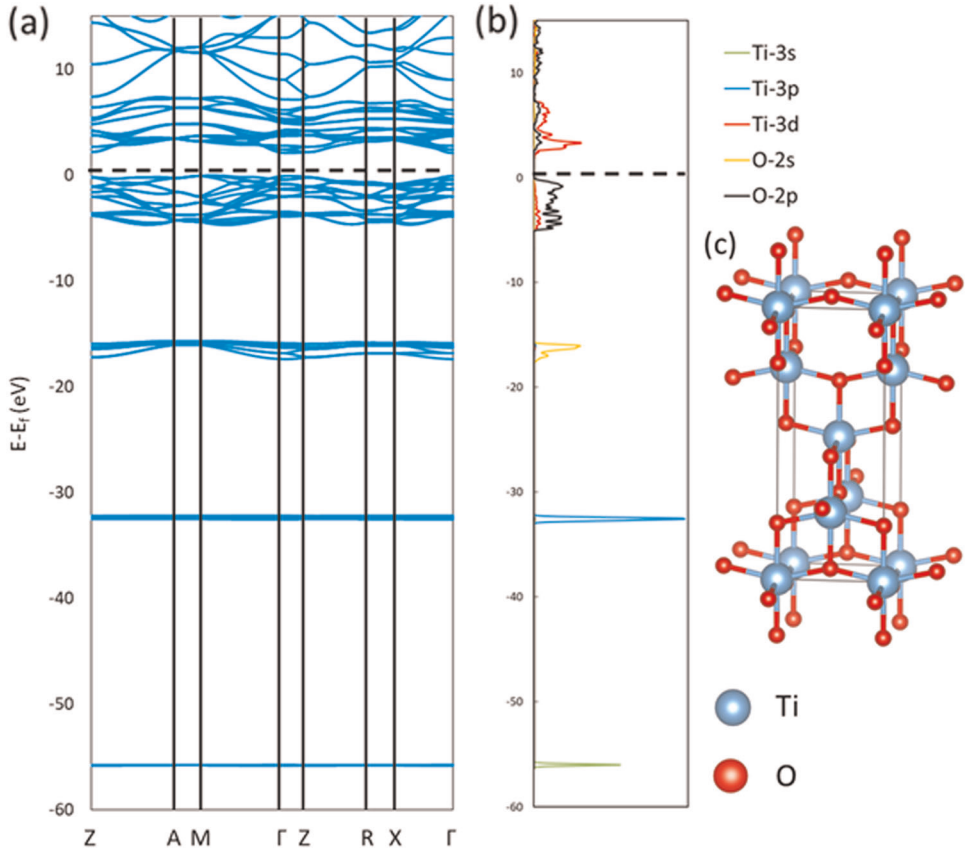

Fig. 2 Electronic structure and atomic structure of anatase TiO_2. a Band structure calculated by SG15 pseudopotential of anatase TiO_2, **b** projected density of states of anatase TiO_2, and **c** the structure of anatase TiO_2.

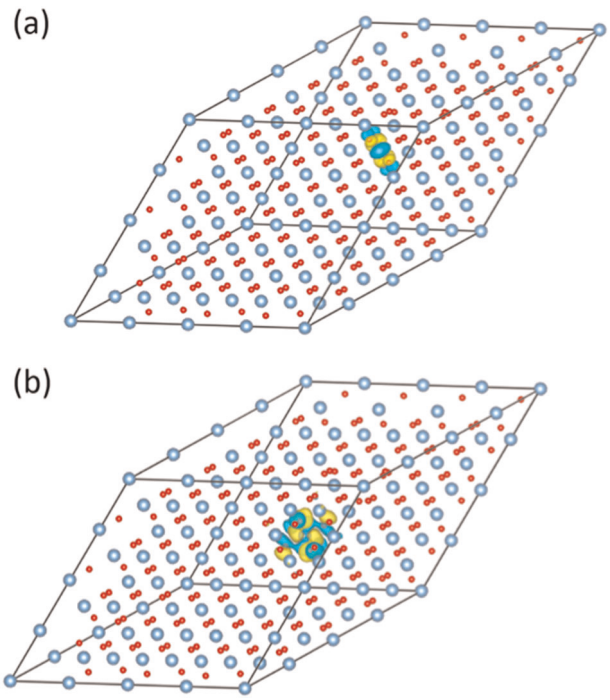

Fig. 3 Most projected wannier functions used in TiO_2 calculations. Wannier functions in TiO_2 system with most projection portion in **a** CBM with projection of 1 and **b** VBM with projection of 0.72. The initial guess of these two Wannier functions are Ti-dxz and O-p, respectively.

Table 1. Calculated anatase TiO_2 WKM band gaps with various numbers of fixed bands.

Fixed bands	CB	VB	LDA	WKM
No fixed bands	1.23	1.34	1.80	4.37
Ti-3s	1.19	1.34	1.80	4.33
Ti-3s, Ti-3p	0.96	1.34	1.80	4.10
Ti-3s, Ti-3p, O-2s	0.65	1.13	1.80	3.58

CB stands for the average λ calculated for conduction band and VB stands for the average λ calculated for valence band.

Table 2. Calculated AlAs WKM band gaps with and without fixed core states.

Fixed bands	CB	VB	LDA	WKM
No fixed bands	0.50	0.67	1.27	2.44
Al-2s	0.50	0.67	1.27	2.44
Al-2s, Al-2p	0.48	0.60	1.27	2.35
Al-2s, Al-2p, As-3s	0.48	0.60	1.27	2.35

CB stands for the average λ calculated for conduction band and VB stands for the average λ calculated for valence band.

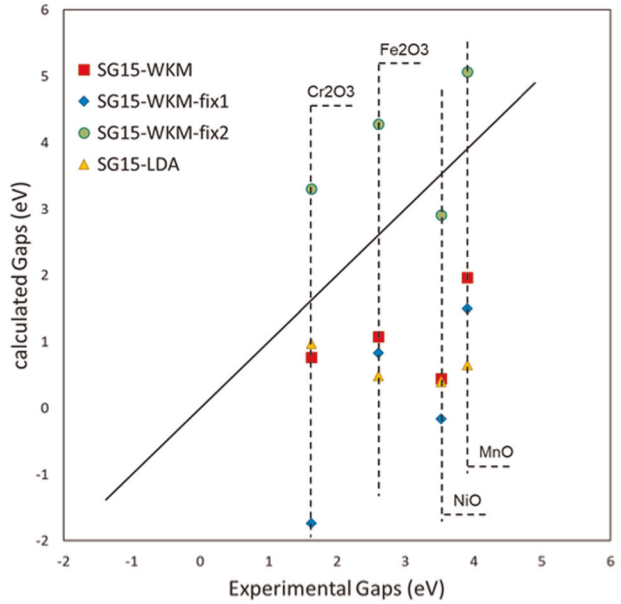

Fig. 4 The calculated WKM band gaps in SG15 pseudopotential. The black line indicates the experimental band gaps. Yellow triangles labeled as SG15-LDA indicate LDA band gaps. All WKM calculations are done with the spin-up and spin-down numbers fixed. Red quadrangles labeled as SG15-WKM indicate WKM results. Blue diamonds labeled as SG15-WKM-fix1 indicate WKM results with only semicore level fixed. Green circles labeled as SG15-WKM-fix2 indicate WKM results with both semicore level and the opposite spin channel (the channel without Wannier function) fixed.

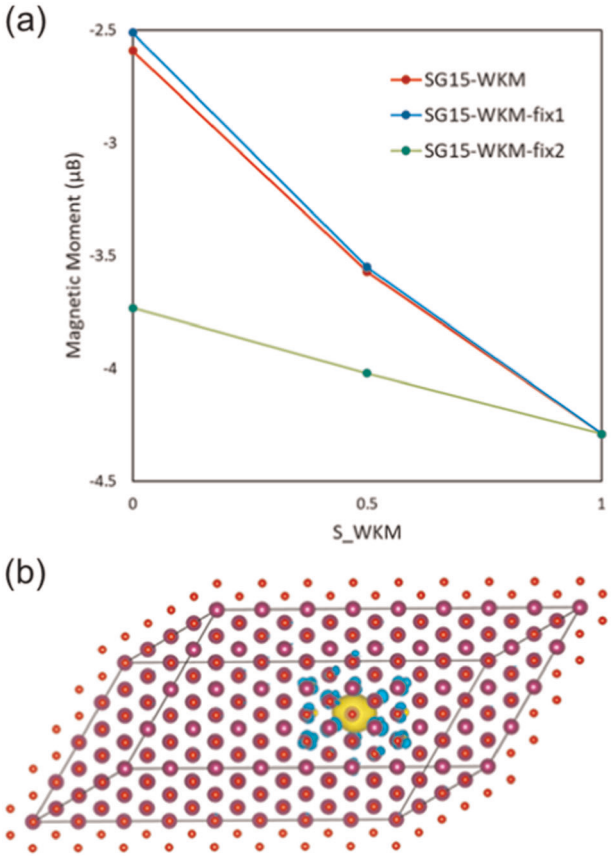

Fig. 5 The changes of charge density in real space in WKM calculations. a The magnetic moment of the Mn atom where d_{x2-y2} Wannier function locates. The red line labeled as SG15-WKM indicate normal WKM calculation without any band fixation. The blue line labeled as SG15-WKM-fix1 indicates SG15-WKM calculation with core-level fixed. The green line labeled as SG15-WKM-fix2 indicates SG15-WKM calculation with core level and a spin channel (the channel without Wannier function) fixed. **b** The charge difference in spin-up channel (the channel without Wannier function) between SG15-WKM-fix2 calculation and a normal self-consistent calculation. The yellow part indicates after an electron of Wannier function was removed from spin-down channel, and the charge density in SG15-WKM-fix2 calculation increased. The blue part indicates the charge density decreased.

part, we tested MnO and NiO in rock-salt structure, and Fe_2O_3 and Cr_2O_3 in corundum structure. All these structures are in antiferromagnetic phase, while previous d^0 and d^{10} TM oxides are in non-magnetic phases. The results are shown in Fig. 4, while the original data can be found in Supplementary Table 8. Because FHI–LDA gaps are not consistent with SG15-LDA gaps for these systems, and their ground-state magnetic moments are also very different (see Supplementary Tables 13 and 14), here we have only considered the SG15 results.

During the $E_{LDA}(\{s_w\})$ calculation of these open shell systems, we found that the number of electrons in a given spin can change dramatically after s_w electron is placed inside one Wannier function. Very often, the loss (or increase) of the charge from one spin will cause a charge transfer between spin channels. This might be caused by a strong on-site interaction between the occupied d-Wannier (constructed from the valence band) and unoccupied d-Wannier (constructed from the conduction band). For example, in the ground state of NiO (an antiferromagnetic system in LDA calculation), the number of electrons in spin-up and spin-down channels are the same (both are 1536 electrons in a supercell). However, after adding one electron to the unoccupied d ϕ_w of spin-down channel in a WKM supercell SCF calculation, the converged ground state will have 1534.56 electrons in the spin-up channel and 1538.44 electrons in the spin-down channel (instead of the expected 1537 electron if there was no spin flow between these two channels). Corresponding to this spin flow, the system lost its band gap, and became metallic (judged by the eigenenergies for orbitals besides the Wannier function). Such spin flow between the spin-up and spin-down channels does not exist in the WKM calculation of d^0 or d^{10} systems, like ZnO.

The above strong on-site interaction between the opposite spin Wannier functions is a manifestation of the strongly correlated system. To avoid this spin flow between opposite spins, for all the $E_{LDA}(\{s_w\})$calculation below, we have forced the number of spin-up and spin-down electrons to be unchanged in the self-consistent iteration (so spin up and spin down will have different Fermi energies).

The tested results are shown in Fig. 4. We have tested the normal WKM procedure (shown as SG15-WKM), and the procedure removing the semicore in the E_{hxc} calculations (shown as SG15-WKM-fix1), both have forced the number of spin-up and spin-down electrons unchanged during SCF calculations. As one can see, although removing semicore procedure does reduce the WKM gap as it did on the d^0 and d^{10} materials, it can make the result even worse when compared with the experiments. For Cr_2O_3, Fe_2O_3, and NiO, the WKM band gaps with fixed/excluded semicore using SG15 are even smaller than that of LDA (see also Supplementary Table 9). Thus, fixing the spin-up and spin-down electron number during WKM SCF calculation does not solve the problem.

The remaining problem can be analyzed by looking at the magnetic moment at each atom during WKM SCF calculations (see Supplementary Tables 10 and 11). Taking MnO as an example, the Mn LDA magnetic moment for SG15 pseudopotential calculation is −4.3 μ_B. In Fig. 5a, we show that, in both SG15-WKM and SG15-WKM-fix1 calculations (where the numbers of electron in the spin-

up and spin-down channels are fixed, so there is no spin flow between these two channels), after removing one electron from the valence d_{x2-y2} Wannier function (shown in Supplementary Fig. 5) in the spin-down channel, the absolute value of magnetic moment for the atom at the center of the Wannier function has changed from -4.3 to $-2.5\ \mu_B$. This reduction is >1 (the removal of the electron from the down spin).

This is caused not by a spin flow between the spin-up and spin-down channels, as their numbers are fixed. Rather, it is caused by a spatial charge flowing within the same spin, mostly from the opposite spin (the up spin). More specifically, the charge in the up spin, from the neighboring Mn atoms from the occupied Wannier function, has flowed to the unoccupied up spin Wannier function at the center. In another word, there is an interaction between the occupied and unoccupied d-state Wannier functions between neighboring atoms at the same spin. This is shown in Fig. 5b, where the spin-up electron charge density after the SCF $E_{\mathrm{LDA}}(\{s_w\})$ calculation increases at the center atom and decreases around the neighboring atoms. Such charge flow is driven by the local total charge (spin up plus spin down) neutrality requirement to minimize the total energy and enabled by the interaction between the conduction and valence band d-state Wannier functions. This is not the case for the Wannier functions of other orbitals. For instance, in NiO, after removing one electron from the O-p_x orbital Wannier function, the magnetic moment changed from 0 to 0.81 and the change is <1.

The above spatial charge flow can also be analyzed directly using Wannier occupation $o_w = \sum_i |\langle \phi_w | \psi_i \rangle|^2 o_i$, instead of atomic magnetic moments. Before the WKM SCF calculations (e.g., using LDA ψ_i), o_w is 0 for conduction band Wannier function ϕ_w, and 1 for valence band Wannier function ϕ_w. If we take one electron from one valence spin-down Wannier function ϕ_w, then the o_w of the neighboring valence spin-up Wannier functions will become less than 1 after SCF calculation, and o_w for the center conduction band spin-up Wannier function will have $o_w > 0$ (e.g., 0.36 for NiO). This corroborates well with the atomic magnetic moment analysis.

To avoid such opposite spin spatial charge flow, we carried out a test where the wave functions as well as their occupations of the opposite spin are fixed at their ground-state LDA values when performing the SCF $E_{\mathrm{LDA}}(\{s_w\})$ calculations. We call this the "fix2" procedure in Figs. 4 and 5. The "fix2" procedure has indeed resulted in smaller than one magnetic moment changes when we remove (or add) one electron from one d-state Wannier functions, as shown in Fig. 5a (see also Supplementary Table 12). Fixing the opposite spin wave function (not just its total number of electron) does increase the band gap, but it overestimates the band gap significantly. This is because by fixing the opposite spin wave functions, we have completely removed their dielectric screening effect, which lead to a too big self-interaction energy for the Wannier function, thus too large λ_w values and band gaps. So far, we have not found a reliable way to calculate the band gap of these Mott insulators under the current WKM approach. Compared with the d^0 and d^{10} systems, one challenge here is the strong interaction and coupling between the two spin channels via their d-state Wannier functions from the opposite sides of the band gap. In a more conventional semiconductor, or in the d^0 and d^{10} cases, such strong interaction does not exist.

The difficulty of these Mott insulators might also be related to the DFT ground states themselves. It could be true that these DFT ground states (e.g., the occupied electrons wave function subspace) are not correct. This can be viewed by their atomic magnetic moment (which represents the amount of d-state occupation). Supplementary Table 14 shows that the DFT calculated magnetic moment can be very different from the experimental value. For instance, the magnetic moment in LDA calculation for NiO is $1.1\ \mu_B$, whereas the experimental value is from 1.64 to $1.9\ \mu_B$. Since our current WKM method does not mix the valence band wave function subspace with conduction band wave function subspace, it is unlikely that it can fix this DFT ground-state problem.

Given the strong coupling between the occupied and unoccupied d-state Wannier functions, and the need to correct the DFT ground-state occupation subspace, it seems like, to find a solution for the partially occupied d-state systems, it might be necessary to use Wannier functions not constructed from occupied and unoccupied orbitals separately, instead d-state Wannier functions constructed from both occupied and unoccupied orbitals might be necessary. Such Wannier functions can be more localized, and the dynamics and coupling between them must be dealt with coherently. The Wannier functions from neighboring atomic sites might also need to be treated together. This is much like the LDA + U approach. For example, Cococcioni et al.[19] have shown that, in an LDA + U method, a linear response approach to calculate U (satisfying the Koopmans requirement) seems to fix the band gap problems in NiO. Such an approach can also change the occupation subspace, thus fixing the DFT ground-state problem.

It is worth to note that, not just WKM has a difficulty in describing such strongly correlated Mott insulator, other mean-field methods, including the GW method, have similar issues. As discussed in the Introduction, the GW results for MnO bandgaps varies from 2.3[13] to 4.4 eV[15] depending on the self-consistent procedure and initial input wave functions. The same is true for Fe_2O_3. For example, Liao and Carter[17] investigated the GW band gap for α-Fe_2O_3. They found that the calculated GW band gap can range from 1.3 to 4.8 eV depending on the initial input (from PBE, LDA + U to HSE06, PBE0), while the experimental band gap is 2.6 eV.

It is also an interesting general question whether the strongly correlated Mott insulator can be described by a Slater determinant in the wave function expression. Although the current WKM scheme fails to describe such system, this does not exclude the possibility that some variance of it, e.g., the one with Wannier function from both occupied and unoccupied orbitals might be successful. It remains to be seen whether one has to use multi-configuration or local correlated treatment like in the dynamic mean-field theory to describe such systems.

In conclusion, we tested the WKM for the TM oxide systems. We found that it can predict well the band gaps for the d^0 and d^{10} TM oxides, provided that a semicore removal procedure is used in the E_{hxc} calculation. On the other hand, the current WKM procedure fails to predict the band gap for the partially occupied d-state TM oxides, even after tests of several procedures. The problem seems to have originated from the strong interaction between the valence band and conduction band d-state Wannier functions. To solve this problem, one has to go beyond the current WKM procedure. Instead, the d-state Wannier functions constructed from a combination of valence band and conduction band orbitals should be used, and their mutual interactions need to be treated coherently.

METHODS

All our calculations are performed using PWmat[41,42], which runs on graphics processing unit. Fifty-Ryd plane wave cutoff was used in our calculations for SG15 pseudopotential. For FHI pseudopotential, the pseudopotential recommended cutoff energies for different elements are used accordingly. Monkhorst–Pack[43] method is used for k-points sampling. In order to ensure that the Wannier functions are orthogonal to semicore states, k-points setting for the bulk (for Wannier function generation) was consistent with our supercells (e.g., the bulk k-points are folded from the supercell Γ point). Generally speaking, the number of k-points multiplied by the number of atoms are >500 and <1000. Our supercell size is thus also between 500 and 1000 atoms, and only Γ point is used for supercell calculations. The calculated LDA band gaps between FHI and SG15 are typically very close as shown in Fig. 1. Wannier90 code[37] was used to generate Wannier functions. Initial guess of Wannier functions is needed

for this step. Azimuthal quantum number and magnetic quantum number of the initial guess are determined by the projected density of states of the system. Usually, the Wannier functions look like their atomic orbitals. For clarity, in the following, a d-orbital means a Wannier function generated by an initial guess of d-orbital. In order to get a localized Wannier function, for d-orbital in octahedrons and tetrahedrons, x, y, z directions were set as the direction, which point to its neighbor oxygen in octahedrons or the cubic high-symmetry directions in tetrahedrons. The directions are shown in Supplementary Fig. 4.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

The WKM code related to this article is implemented in PWmat commercial code with official web page http://www.pwmat.com/.

Received: 18 September 2019; Accepted: 10 March 2020;
Published online: 03 April 2020

REFERENCES

1. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965).
2. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
3. Zhao, Y. & Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **125**, 194101 (2006).
4. Heyd, J., Scuseria, G. E. & Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **118**, 8207–8215 (2003).
5. Kim, K. & Jordan, K. D. Comparison of density functional and MP2 calculations on the water monomer and dimer. *J. Phys. Chem.* **98**, 10089–10094 (1994).
6. Stephens, P. J., Devlin, F. J., Chabalowski, C. F. & Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **98**, 11623–11627 (1994).
7. Anisimov, V. I. & Gunnarsson, O. Density-functional calculation of effective Coulomb interactions in metals. *Phys. Rev. B* **43**, 7570–7574 (1991).
8. Hedin, L. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* **139**, A796–A823 (1965).
9. Strehlow, W. H. & Cook, E. L. Compilation of energy band gaps in elemental and binary compound semiconductors and insulators. *J. Phys. Chem. Ref. Data* **2**, 163–200 (1973).
10. Samsonidze, G., Park, C. H. & Kozinsky, B. Insights and challenges of applying the GW method to transition metal oxides. *J. Phys. Condens. Matter* **26**, 475501 (2014).
11. Van Elp, J., Potze, R. H., Eskes, H., Berger, R. & Sawatzky, G. A. Electronic structure of MnO. *Phys. Rev. B* **44**, 1530–1537 (1991).
12. Faleev, S. V., van Schilfgaarde, M. & Kotani, T. All-electron self-consistent approximation: application to Si, MnO, and NiO. *Phys. Rev. Lett.* **93**, 126406 (2004).
13. Jiang, H., Gomez-Abal, R. I., Rinke, P. & Scheffler, M. First-principles modeling of localized d states with the GW@LDA+U approach. *Phys. Rev. B* **82**, 1–16 (2010).
14. Lany, S. Band-structure calculations for the 3d transition metal oxides in GW. *Phys. Rev. B* **87**, 1–9 (2013).
15. Das, S., Coulter, J. E. & Manousakis, E. Convergence of quasiparticle self-consistent calculations of transition-metal monoxides. *Phys. Rev. B* **91**, 115105 (2015).
16. Jiang, H. Revisiting the GW approach to d- and f-electron oxides. *Phys. Rev. B* **97**, 1–9 (2018).
17. Liao, P. & Carter, E. A. Testing variations of the GW approximation on strongly correlated transition metal oxides: hematite (α-Fe2O3) as a benchmark. *Phys. Chem. Chem. Phys.* **13**, 15189–15199 (2011).
18. Dabo, I. et al. Koopmans' condition for density-functional theory. *Phys. Rev. B* **82**, 115121 (2010).
19. Matteo Cococcioni, S. & de, G. Linear response approach to the calculation of the effective interaction parameters in the LDA+U method. *Phys. Rev. B* **71**, 35105 (2005).
20. Kraisler, E. & Kronik, L. Piecewise linearity of approximate density functionals revisited: implications for frontier orbital energies. *Phys. Rev. Lett.* **110**, 126403 (2013).
21. Ma, J. & Wang, L.-W. Using Wannier functions to improve solid band gap predictions in density functional theory. *Sci. Rep.* **6**, 24924 (2016).
22. Weng, M. et al. Wannier Koopman method calculations of the band gaps of alkali halides. *Appl. Phys. Lett.* **111**, 054101 (2017).
23. Weng, M., Li, S., Zheng, J., Pan, F. & Wang, L.-W. Wannier Koopmans method calculations of 2D material band gaps. *J. Phys. Chem. Lett.* **9**, 281–285 (2018).
24. Li, S. et al. Wannier-Koopmans method calculations of organic molecule crystal band gaps. *Europhys. Lett.* **123**, 37002-p1–37002-p6 (2018).
25. Li, C., Zheng, X., Su, N. Q. & Yang, W. Localized orbital scaling correction for systematic elimination of delocalization error in density functional approximations. *Natl Sci. Rev.* **5**, 203–215 (2018).
26. Nguyen, N. L., Colonna, N., Ferretti, A. & Marzari, N. Koopmans-compliant spectral functionals for extended systems. *Phys. Rev. X* **8**, 1–9 (2018).
27. Zhan, C.-G., Nichols, J. A. & Dixon, D. A. Ionization potential, electron affinity, electronegativity, hardness, and electron excitation energy: molecular properties from density functional theory orbital energies. *J. Phys. Chem. A* **107**, 4184–4195 (2003).
28. Janak, J. F. Proof that dE/dn_i = e_i in density-functional theory. *Phys. Rev. B* **18**, 7165–7168 (1978).
29. Jain, A. et al. Commentary: The Materials Project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
30. Bockstedte, M., Kley, A., Neugebauer, J. & Scheffler, M. Density-functional theory calculations for poly-atomic systems: electronic structure, static and elastic properties and ab initio molecular dynamics. *Comput. Phys. Commun.* **107**, 187–222 (1997).
31. Hamann, D. R. Generalized norm-conserving pseudopotentials. *Phys. Rev. B* **40**, 2980–2987 (1989).
32. Fuchs, M. & Scheffler, M. Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory. *Comput. Phys. Commun.* **119**, 67–98 (1999).
33. Fartash, A., Schuller, I. K. & Grimsditch, M. Thin-film modeling for mechanical measurements: should membranes be used or plates? *J. Appl. Phys.* **71**, 4244–4248 (1992).
34. Troullier, N. & Martins, J. L. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **43**, 1993–2006 (1991).
35. Hamann, D. R. Optimized norm-conserving Vanderbilt pseudopotentials. *Phys. Rev. B - Condens. Matter Mater. Phys.* **88**, 1–10 (2013).
36. Schlipf, M. & Gygi, F. Optimization algorithm for the generation of ONCV pseudopotentials. *Comput. Phys. Commun.* **196**, 36–44 (2015).
37. Souza, I., Marzari, N. & Vanderbilt, D. Maximally localized Wannier functions for entangled energy bands. *Phys. Rev. B* **65**, 035109 (2001).
38. Dette, C. et al. TiO2 anatase with a bandgap in the visible region. *Nano Lett.* **14**, 6533–6538 (2014).
39. Lorenz, M. R., Chicotka, R., Pettit, G. D. & Dean, P. J. The fundamental absorption edge of AlAs and AlP. *Solid State Commun.* **8**, 693–697 (1970).
40. Antonius, G. et al. Dynamical and anharmonic effects on the electron-phonon coupling and the zero-point renormalization of the electronic structure. *Phys. Rev. B* **92**, 1–9 (2015).
41. Jia, W. et al. Fast plane wave density functional theory molecular dynamics calculations on multi-GPU machines. *J. Comput. Phys.* **251**, 102–115 (2013).
42. Jia, W. et al. The analysis of a plane wave pseudopotential density functional theory code on a GPU machine. *Comput. Phys. Commun.* **184**, 9–18 (2013).
43. Pack, J. D. & Monkhorst, H. J. 'Special points for Brillouin-zone integrations'—a reply. *Phys. Rev. B* **16**, 1748–1749 (1977).

ACKNOWLEDGEMENTS

L.-W.W. is supported by the Director, Office of Science, the Office of Basic Energy Sciences (BES), Materials Sciences and Engineering (MSE) Division of the U.S. Department of Energy (DOE) through the theory of material (KC2301) program under Contract No. DEAC02-05CH11231. This work is also financially supported by National Materials Genome Project of China (2016YFB0700600) and Shenzhen Science and Technology Research Grant (Nos JCYJ20160226105838578 and JCYJ20151015162256516).

AUTHOR CONTRIBUTIONS

M.W. and F.P. performed WKM calculations, collected the data, and analysed the data. L.-W.W. designed this research, came up with the core-level fixation idea, and led all the discussion and data analysis.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41524-020-0302-0.

Correspondence and requests for materials should be addressed to F.P. or L.-W.W.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.