LETTER

Wannier-Koopmans method calculations of organic molecule crystal band gaps

A Letters Journal Exploring the Frontiers of Physics

August 2018

EPL, **123** (2018) 37002
doi: 10.1209/0295-5075/123/37002

www.epljournal.org

Wannier-Koopmans method calculations of organic molecule crystal band gaps

Shucheng Li[1], Mouyi Weng[1], Jianshu Jie[1], Jiaxin Zheng[1], Feng Pan[1(a)] and Lin-Wang Wang[2(b)]

[1] *School of Advanced Materials, Peking University, Shenzhen Graduate School - Shenzhen 518055, PRC*
[2] *Materials Science Division, Lawrence Berkeley National Laboratory - Berkeley, CA 94720, USA*

received 6 May 2018; accepted in final form 31 July 2018
published online 4 September 2018

PACS 71.15.-m – Methods of electronic structure calculations

Abstract – It is important to accurately predict the band gaps of crystals, including organic crystals, with low computational cost. Despite the significant underestimation of the crystal band gap by the density functional theory (DFT), a recently proposed Wannier-Koopmans method (WKM) based on DFT calculations seems to yield accurate band gaps for a wide class of materials including common semiconductors, alkali halides and 2D materials. It is nevertheless important to test the limit of WKM, in particular in systems with unique characteristics. In this work, we apply the WKM to 10 organic small molecule crystals and find that the WKM calculated band gaps agree well with GW results. We also introduce a new way to calculate the Wannier functions in the WKM calculations.

Introduction. – Small organic molecule crystals are a class of very interesting material due to their flexibility, low-temperature processability, light weight and low cost. They have been used in organic field-effect transistors (OFETs) [1], light-emitting diodes (OLEDs) [2] and photovoltaic cells (OPVs) [3]. For many of these applications, the quasiparticle band gap is a critical device parameter one wishes to calculate from *ab initio* methods. Unfortunately, it is well known that the widely used density functional theory (DFT) [4] tends to significantly underestimate the band gap [5]. Various approaches have been tested to go beyond the local density approximation (LDA) [6] and generalized gradient approximation (GGA) [7] in order to correct the band gap. For d and f orbital systems, the DFT$+U$ method has been used to partially remove the self-interaction energy [8]. Although it does open up the band gaps for many transition metal oxides, the parameter U is fitted to experiments [9], thus it loses its power to predict new systems in an *ab initio* fashion. The self-interaction correction (SIC) method [10] explicitly removes the self-interaction term, but for many common semiconductors, the self-consistently solved orbital is not localized, resulting in the same band gap as the original Kohn-Sham method. There are also meta-GGA approaches [11], which use the kinetic energy density to construct more complicated exchange-correlation energy, and for some cases [12] it does increases the energy band gap. But their wide applicability is yet to be proved. There are also hybrid functionals, like HSE [13] and B3LYP [14,15] which improve the band gaps. But they are parameter based, which can lead to inaccurate results, as will be shown in this study, if the same parameters are used for different systems. So far the most accurate and reliable approach for band gap calculations is the GW method [16] based on the many-body perturbation theory. The GW methods have several versions of self-consistency such as G_0W_0, GW_0 and GW. For many systems, the single-shot G_0W_0 calculation often slightly underestimates the experimental value [12], while the GW_0 seems to yield more accurate results [17,18], and the full self-consistent GW calculations can overestimate the band gap [19]. Overall, the GW method can be expensive, and their formalism which involves the high-frequency dynamic dielectric screening and self-energy is very different from the ground-state DFT formalism. It is thus still interesting to develop a method which is more based on the DFT formalism, at the same time also to correct the DFT band gap error.

Recently, we have developed a Wannier-Koopmans method (WKM) [20] to correct the DFT band gap. The WKM extends the well-known ΔDFT method [21], which works well for isolated systems, to extended bulk systems. In the ΔDFT method, one calculates the electron affinity (EA) as $E(N+1)-E(N)$ and ionization energy (IE) as

(a) E-mail: panfeng@pkusz.edu.cn (corresponding author)
(b) E-mail: lwwang@ibl.gov (corresponding author)

$E(N) - E(N-1)$ using self-consistent ground-state energies $E(N)$, $E(N+1)$, and $E(N-1)$, where N, $N+1$, $N-1$ are the corresponding number of electrons in the system. Despite the great success for atoms and small isolated molecules [21], the ΔDFT method fails for extended bulk. In an extended system, the change of charge density by adding one electron in the system is infinitely small at any given spatial point, as a result, the total energy difference is the same as the Kohn-Sham orbital energy due to Janak's theory [22]. To overcome this problem, the WKM method places the additional electron in the localized Wannier functions instead of the extended Kohn-Sham orbitals, so the charge density change is not infinitely small, and a finite correction can be obtained. More specifically, we can express the total energy in the WKM formalism as

$$E_{WKM} = E_{LDA}(\{s_w\}) + \sum\nolimits_w E_w(s_w), \tag{1}$$

where "w" denotes different orthogonal Wannier functions in either occupied manifold or unoccupied manifold and $0 < s_w < 1$ stands for the occupation number of the corresponding Wannier function w. To satisfy the Koopmans theory, the above total energy E_{WKM} is made to be linear with the Wannier function occupation s_w. This is done by of the $E_w(s_w)$ term, which can be approximated as

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

The λ_w is obtained by first calculating $E_{LDA}(\{S_w\})$, which is the total LDA energy with the s_w electron added in (subtracted from) the conduction (valence) band Wannier function ϕ_w. Then, $\lambda_w = \mathrm{d}^2 E_{LDA}(\{s_w\})/\mathrm{d}s_w^2|_{s_w=0/1}$ (0 for conduction band, 1 for valence band). For a given system, s_w can also be calculated as $s_w = \sum_i \langle\phi_w|\psi_i\rangle^2 o(i)$, where ψ_i is the i-th the Kohn-Sham orbital, $o(i)$ is the occupation of ψ_i. As a result, when we seek the variational minimum of E_{WKM} with regard to ψ_i, we yield a modified Kohn-Sham equation:

$$[H_{LDA} \pm \sum\nolimits_w \lambda_w |\phi_w\rangle\langle\phi_w|]\psi_i = \varepsilon_i \psi_i. \tag{3}$$

Here $+$ is for conduction band, $-$ is for valence band. In an actual calculation, the neutral system ground state with N electrons is first calculated. The Kohn-Sham orbitals ψ_i are then used to construct the Wannier function ϕ_w represented in a supercell. $E_{LDA}(\{S_w\})$ is then calculated using this supercell, with ψ_i in the same spin of the occupied ϕ_w being orthogonal to ϕ_w. The variational minimum of $E_{LDA}(\{S_w\})$ with regard to ψ_i under this constraint is self-consistently calculated. This self-consistent solution describes the screening of the Coulomb potential caused by the occupation of ϕ_w. The details of the calculation were described in ref. [23].

In our recent works, the performances of WKM in common semiconductors [23], alkali halides [24] and 2D materials [25] have been tested, so is its ability to describe band alignments [26]. In these systems, WKM has achieved band gap accuracy comparable to *GW* results, and all agree well with experiments. The typical band gap error compared with experiment is in a few tenths of eV [23–25]. Given all these previous works, it is thus extremely interesting to know what the accuracy is of WKM to predict the band gaps of other systems. In this work, we use WKM to calculate the band gaps of 10 organic molecules. Due to the lack of experimental results for their quasiparticle band gaps, we have compared our WKM results with GW results, as well as the GGA and HSE results.

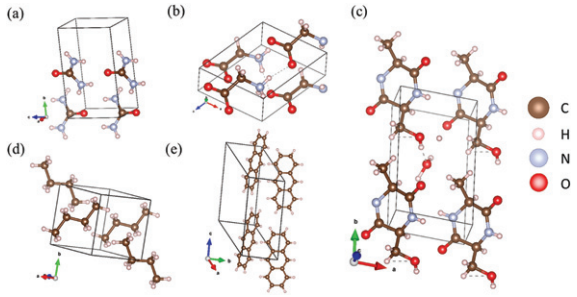

Fig. 1: (Colour online) The structures of the organic molecule crystal. (a), (b), (c), (d), (e) are, respectively, represented as Carbamide, Glycine, (3S,6S)-3-(Hydroxymethyl)-6-methyl-2,5-piperazinedione hydrate, *n*-Butane, Anthracene.

Besides testing this important new class of systems, we have also introduced a new way of generating the Wannier functions. In our previous work, the Wannier90 program was used to calculate the Wannier function. This requires the input of atomic characteristic of the Wannier function. Although it is easy to do this for small inorganic crystals as in our previous works, for large unit cell as for the organic crystals, that becomes untrivial because there could be many different choices for the Wannier function generation. In the current work, we use the selected columns of the density matrix (SCFM) method [27] to generate the Wannier function. This method does not need the prior selection of the atomic character of the Wannier function, which makes the Wannier function generation more automatic.

Computational method. – The selected molecules cover a range of structure types. The simplest molecule is Carbamide and Glycine, while the complex ones include (3S,6S)-3-(Hydroxymethyl)-6-methyl-2,5-piperazinedione hydrate. We have also selected aliphatic molecules like normal butane as well as aromatic molecules like Anthracene. These five crystal structures are shown in fig. 1. We also cover different functional groups, like carboxyl group, nitrile group, ester group, hydroxyl group as shown in fig. S1 of the Supporting Information `Supplementarymaterial.pdf` (SI). All the crystal structures are shown in the SI, and we have used the crystal lattice constants, internal molecule configurations and positions from the Crystallography Open Database [28].

In our calculation, we first used the DFT method to calculate the neutral system ground-state wave function ψ_i.

Table 1: The results calculated by different Wannier functions compared with experimental band gap values of the AlAs system.

	CB (eV)	VB (eV)	LDA (eV)	WKM (eV)	Exp (eV)
Wannier90	0.308	0.688	1.302	2.298	2.220
SCDM-k	0.339	0.698	1.302	2.338	2.220

Table 2: The results calculated by different Wannier functions compared with GW calculated band gap values of the Glycine system.

	CB (eV)	VB (eV)	LDA (eV)	WKM (eV)	GW_0 (eV)
Wannier90	1.473	1.941	5.331	8.745	9.521
SCDM-k	1.459	1.946	5.331	8.736	9.521

This is a primary cell calculation with multiple k-points. These Bloch states are then used in the following Wannier function generation step. In this step, we have used a newly developed and highly parallelizable SCDM-k [27] code, and the results will be compared with the more common Wannier90 code [29].

To test the results between the SCDM-k method and the Wannier90 method for the Wannier function generation, we have used AlAs and Glycine systems as examples. For the AlAs system we have obtained similar band gaps using Wannier90 and SCDM-k, 2.30 eV for Wannier90 and 2.34 eV for SCDM-k, respectively, while the experimental value is 2.22 eV as shown in table 1. For the Glycine system, we did not find the experimental quasiparticle gap, so we have compared the band gap with the GW_0 result shown in table 2. The band gaps calculated by Wannier90 and SCDM-k are 8.745 eV and 8.736 eV, both of them are within a 10% of the GW_0 band gap (9.521 eV). We have compared the Wannier functions generated by the Wannier90 code and SCDM-k code for the Glycine system. As shown in fig. 2, the Wannier functions with the maximum projection to the conduction band minimum generated by the SCDM-k code and Wannier90 are rather similar. The dot product of these two Wannier functions: $\langle\phi_{Wannier90}|\phi_{SCDM\text{-}k}\rangle$ is 0.9983, indicating high similarity. The same similarity between the SCDM-k and Wannier90 results exists for the first few Wannier functions. We have also analyzed the localization of the Wannier functions with maximum band edge state projection. The charge accumulation function from the center of the Wannier function can be calculated as

$$Q(r) = \int_{|r'|<r} w(r')\mathrm{d}^3r'.$$

The charge density integral is plotted in fig. 2(c) for a $3 \times 3 \times 3$ unit super cell. From fig. 2(c) we estimate that the radius of this Wannier function is about 2.84 Å and the minimum distance between two imaging Wannier functions is around 9.67 Å. The minimum distance between imaging Wannier functions is larger than the one used in our previous works [25], hence the supercell size should be sufficient to calculate λ_w. The above analysis establishes the procedure to use WKM to calculate the organic small molecule crystals using the SCDM-k code as the Wannier function generator.

In this work, all DFT calculations were performed using the PWmat code [30,31], which runs on graphics processing unit processors (GPU). NCPP-SG15-PBE pseudopotential [32,33] and 50 Ryd plane wave cutoff was used in all of our calculations, and the k-points were generated by using the Monkhorst-Pack method [34]. The number of k-points multiplied by the number of atoms is larger than 2000 in all of our WKM calculations. In our procedure, we first generate Kohn-Sham wave functions in the PWmat code, which are then used as the input to generate the Wannier functions using the SCDM-k code. We have chosen sufficient sizes of the supercell which ensure that the distance between two imaging Wannier functions is larger than 6.5 Å as shown in fig. 2(c). Lastly, we have calculated $E_{LDA}(\{s_w\})$ of the Wannier function ϕ_w using the PWmat code for Wannier functions with the largest projections on conduction band minimum (CBM) or valence band maximum (VBM) state. Typically, 4–5 Wannier functions are used in order for the total projections on CBM and VBM to be larger than 90%. Three s_w values are calculated for $E_{LDA}(\{s_w\})$ for each ϕ_w. These three values of $E_{LDA}(\{s_w\})$ are used to calculate its second derivative, hence the value of λ_w. The CBM and VBM eigen energy corrections are then provided by eq. (3) using the energy expectation value based on the ground-state Kohn-Sham wave function ψ_i with the value of the Wannier function projection $\langle\phi_w|\psi_i\rangle$ provided by the SCDM-k code. The corrections and the final band gaps of WKM are shown in table S1 of the SI. We have also carried out calculations using G_0W_0 and GW_0 for all ten systems. The G_0W_0 and GW_0 calculations are both implemented in the Vienna *ab initio* simulation package (VASP) [35] code using the projector augmented-wave potential [36,37]. We have used an energy cutoff of 400 eV. We set the number of virtual orbitals for over 100 bands per atom, and we have tested the convergence of the parameter "NBANDS" for GW_0 calculations which are shown in table S4 of the SI. In addition, we have calculated PBE [38] and HSE06 (with HSE parameter alpha 0.25 and omega 0.20) band gaps using the PWmat code. Lastly, the LDA band gaps, as the starting point of our WKM method, are also calculated. The computation time of the main methods are listed in table S5 of the SI, although caution must be used in comparing their times since different computers are used and the times sensitively depend on the computational parameters like the k-points. In our calculations, GW_0 and G_0W_0 are calculated on the CPU compute nodes, and LDA, PBE, HSE and WKM are calculated on one

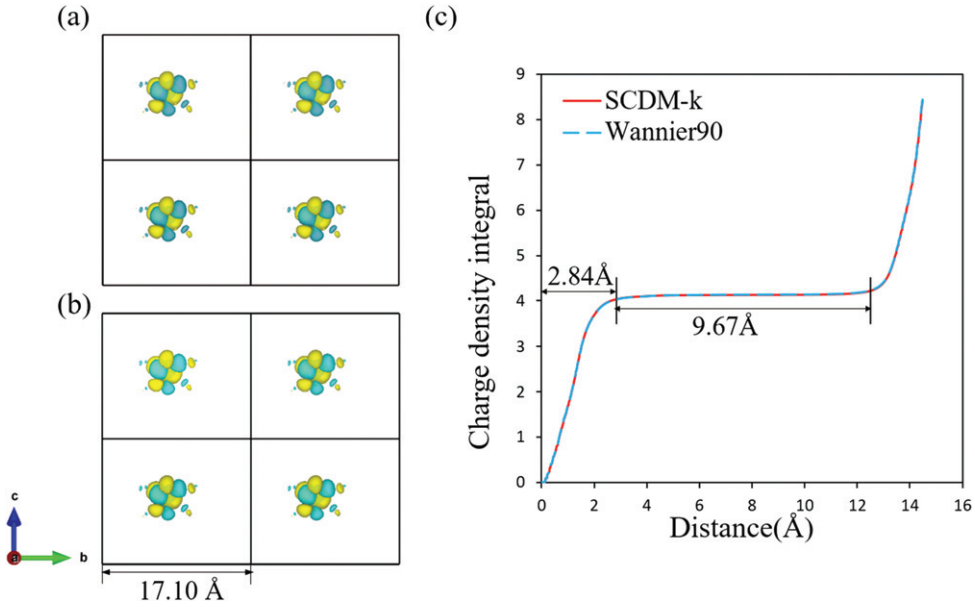

Fig. 2: (Colour online) (a) and (b) are the maximum projection Wannier function of Glycine calculated by SCDM-k and Wannier90. (c) The charge density by the distance from the center of Glycine Wannier functions (a), (b) in a supercell of $3 \times 3 \times 3$.

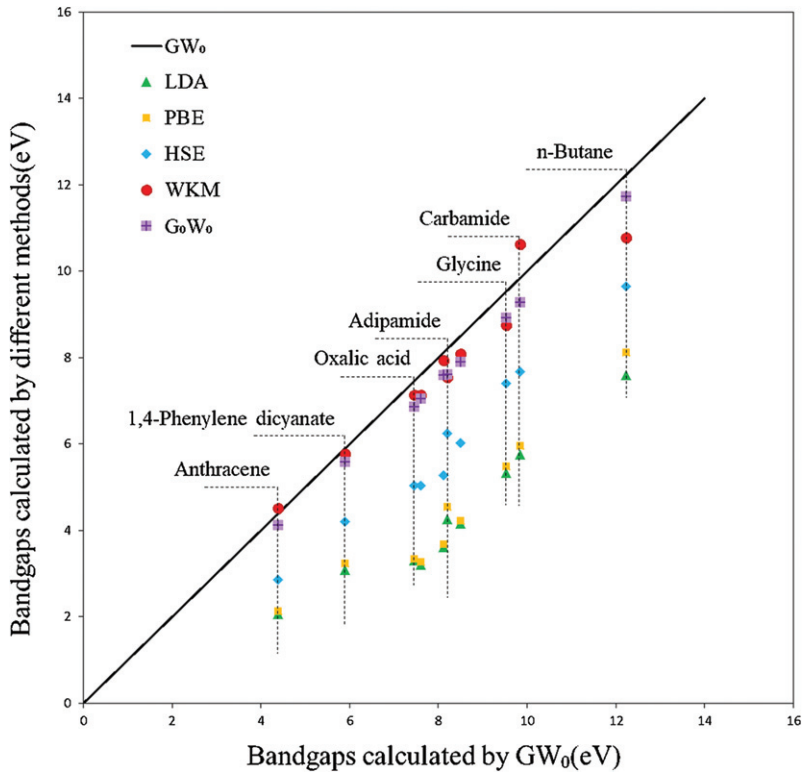

Fig. 3: (Colour online) The calculated WKM band gap, GW_0 band gap, G_0W_0 band gap, LDA band gap, PBE band gap, HSE band gap. The complete result is shown in table S3 of the SI.

GPU node. One CPU node consists of two Intel ® Xeon ® CPU (E5-2640 v4) CPUs where one CPU contains 20 cores, a GPU node consists of four NVIDA ® GTX 1080Ti GPUs. All the above parameters settings are summarized in the SI.

Results and discussions. – All of our calculated results are plotted in fig. 3. Overall, our results agree well with GW_0 and G_0W_0. In general, G_0W_0 results are smaller than the GW_0 results, and in many cases, the WKM results are between the GW_0 and G_0W_0 results.

We also see that the HSE06 significantly underestimates the band gap. This occurs using the HSE06 parameters which yield good results for common semiconductors. We found that it is generally true for larger band gap systems, the HSE06 underestimates the band gap. This is probably because for larger band gap systems, the screening effects are smaller. But the HSE06 parameters are heavily fitted for smaller band gap systems, which have larger screening, hence will have too much screening for large band gap materials, which leads to an underestimation of the band gaps. Note, in previous works [23], we have shown that the WKM also works for isolated molecules, or when an oligomer gradually increases its length becoming a one-dimensional crystal. The applicability of WKM to small molecule crystal implies that WKM is generally applicable for all organic systems. For disorder organic systems (*e.g.*, a polymer blend), the challenge is to calculate all the Wannier functions. One approach is to evaluate the λ_w and VBM, CBM to the Wannier function projection $\langle\phi_w|\psi_i\rangle$ with some approximations, much like in the charge patching method for such disordered systems [39]. In table S6 of the SI, we have listed the individual VBM and CBM corrections under the WKM and GW_0 methods. These individual corrections (instead of the band gap) are important for applications like band alignment. As we can see, not only our band gap agrees with the GW_0 results, the individual VBM and CBN corrections also agree with the GW results. In fig. S2 of the SI, we presented the LDA band structure of the systems. As can be seen, due to the weak molecule-molecule coupling, the dispersions of these band structures are rather small, especially for valence bands. As a result, it is rather difficult to determine the exact VBM k-point locations. Furthermore, not exact k-point grids are used in WKM and GW_0 calculations. Nevetheless, we found that all the band gaps are indirect, and the CBM states are located in the same vicinity within the Brillouin zone between WKM and GW_0 methods.

We also notice that there are small differences between LDA and PBE band gaps, although qualitatively they are the same. Their difference however is in the same order for the difference between WKM and GW_0. This could mean that the remaining error in WKM (*e.g.*, compared to GW_0, or experiment) is in the same level as the difference between LDA and PBE, *i.e.*, the difference between different exchange-correlation (EC) functionals. This is because in WKM, we have removed the self-interaction error in the common EC functionals. After the removal of this main error (*e.g.*, this will introduce the proper discontinuity in the derivative of the EC functional [23]), there could still be other errors (we still do not have the exact EC functional, even for the continuous part of the EC functional). These "other errors" (*e.g.*, due to the continuous part) will equally affect the ground-state property and the excited state (band gap) property. But at least, we will no longer have a "particular problem" for the band gaps for the DFT calculations.

Conclusions. – In conclusion, we have applied our Wannier-Koopmans method to 10 organic small molecular crystals, and used a new SCDM-k method to generate Wannier functions. We found that the WKM gives a similar band gap to that of the GW_0 method. This shows that the WKM is applicable to extended organic systems.

* * *

L-WW is supported by the Director, Office of Science (SC), Basic Energy Science (BES), Materials Science and Engineering Division (MSED), of the U.S. Department of Energy (DOE) under Contract No. DE-AC02-05CH11231 through the Materials Theory program (KC2301). This work is also financially supported by National Materials Genome Project of China (2016YFB0700600), and Shenzhen Science and Technology Research Grant (No. JCYJ20150729111733470 and No. JCYJ20151015162256516).

REFERENCES

[1] Allard S., Forster M., Souharce B., Thiem H. and Scherf U., *Angew. Chem. Int. Ed.*, **47** (2008) 4070.
[2] Lamansky S., Djurovich P., Murphy D., Abdel-Razzaq F., Lee H. E., Adachi C., Burrows P. E., Forrest S. R. and Thompson M. E., *J. Am. Chem. Soc.*, **123** (2001) 4304.
[3] Gregg B. A. and Hanna M. C., *J. Appl. Phys.*, **93** (2001) 3605.
[4] Rajagopal A. K. and Callaway J., *Phys. Rev. B*, **7** (1973) 1912.
[5] Perdew J. P. and Levy M., *Phys. Rev. Lett.*, **51** (1983) 1884.
[6] Kohn W. and Sham L. J., *Phys. Rev.*, **140** (1965) 1133.
[7] Kresse G. and Furthmüller J., *Comput. Mater. Sci.*, **6** (1996) 15.
[8] Anisimov V. I., Zaanen J. and Andersen O. K., *Phys. Rev. B*, **44** (1991) 943.
[9] Wang L., Maxisch T. and Ceder G., *Phys. Rev. B - Condens. Matter Mater. Phys.*, **73** (2006) 1.
[10] Perdew J. P. and Zunger A., *Phys. Rev. B*, **23** (1981) 5048.
[11] Zhao Y. and Truhlar D. G., *J. Chem. Phys.*, **125** (2006) 194101.
[12] Shepard S. and Smeu M., *Comput. Mater. Sci.*, **143** (2018) 301.
[13] Heyd J., Scuseria G. E. and Ernzerhof M., *J. Chem. Phys.*, **118** (2003) 8207.
[14] Kim K. and Jordan K. D., *J. Phys. Chem.*, **98** (1994) 10089.
[15] Stephens P. J., Devlin F. J., Chabalowski C. F. and Frisch M. J., *J. Phys. Chem.*, **98** (1994) 11623.
[16] Hybertsen M. S. and Louie S. G., *Phys. Rev. Lett.*, **55** (1985) 1418.
[17] Shishkin M. and Kresse G., *Phys. Rev. B*, **75** (2007) 1.
[18] Barth U. von and Holm B., *Phys. Rev. B*, **54** (1996) 8411.
[19] Cao H., Yu Z., Lu P. and Wang L. W., *Phys. Rev. B*, **95** (2017) 1.
[20] Ma J. and Wang L. W., *Sci. Rep.*, **6** (2016) 24924.

[21] Zhan C. G., Nichols. J. A. and Dixon D. A., *J. Phys. Chem. A*, **107** (2003) 4184.
[22] Janak J. F., *Phys. Rev. B*, **18** (1978) 7165.
[23] Ma J. and Wang L. W., *Sci. Rep.*, **6** (2016) 1.
[24] Weng M., Li S., Ma J., Zheng J., Pan F. and Wang L. W., *Appl. Phys. Lett.*, **111** (2017) 054101.
[25] Weng M., Li S., Zheng J., Pan F. and Wang L-W., *J. Phys. Chem. Lett.*, **9** (2018) 281.
[26] Ma J., Liu Z. F., Neaton J. B. and Wang L. W., *Appl. Phys. Lett.*, **108** (2016) 262104.
[27] Damle A., Lin L. and Ying L., *J. Comput. Phys.*, **334** (2017) 1.
[28] Graulis S., Chateigner D., Downs R. T., Yokochi A. F. T., Quirós M., Lutterotti L., Manakova E., Butkus J., Moeck P. and Le Bail A., *J. Appl. Crystallogr.*, **42** (2009) 726.
[29] Souza I., Marzari N. and Vanderbilt D., *Phys. Rev. B*, **65** (2001) 35109.
[30] Jia W., Cao Z., Wang L., Fu J., Chi X., Gao W. and Wang L. W., *Comput. Phys. Commun.*, **184** (2013) 9.
[31] Jia W., Fu J., Cao Z., Wang L., Chi X., Gao W. and Wang L. W., *J. Comput. Phys.*, **251** (2013) 102.
[32] Schlipf M. and Gygi F., *Comput. Phys. Commun.*, **196** (2015) 36.
[33] Hamann D. R., *Phys. Rev. B*, **88** (2013) 1.
[34] Pack J. D. and Monkhorst H. J., *Phys. Rev. B*, **16** (1977) 1748.
[35] Kresse G. and Furthmüller J., *Phys. Rev. B*, **54** (1996) 11169.
[36] Blöchl P. E., *Phys. Rev. B*, **50** (1994) 7953.
[37] Joubert D., *Phys. Rev. B*, **59** (1999) 1758.
[38] Perdew J. P., Burke K. and Ernzerhof M., *Phys. Rev. Lett.*, **77** (1996) 3865.
[39] Vukmirović N. and Wang L., *J. Chem. Phys.*, **128** (2008) 121102.