



# AMP<sup>2</sup>: A fully automated program for *ab initio* calculations of crystalline materials<sup>☆,☆☆</sup>

Yong Youn<sup>a</sup>, Miso Lee<sup>a</sup>, Changho Hong<sup>a</sup>, Doyeon Kim<sup>a</sup>, Sangtae Kim<sup>a</sup>, Jisu Jung<sup>a</sup>, Kanghoon Yim<sup>b</sup>, Seungwu Han<sup>a,\*</sup>

<sup>a</sup> Department of Materials Science and Engineering, Seoul National University, Seoul 08826, Republic of Korea

<sup>b</sup> Korea Institute of Energy Research, Daejeon 34129, Republic of Korea

## ARTICLE INFO

### Article history:

Received 6 October 2019

Received in revised form 14 May 2020

Accepted 25 May 2020

Available online 24 June 2020

### Keywords:

Density functional theory

Automation

Crystalline properties

## ABSTRACT

*Ab initio* calculations based on the density functional theory (DFT) become a vital tool in material science for understanding and predicting material properties. However, DFT calculations involve several parameters and procedures that call for deep understanding on underlying theories and preceding knowledge on certain properties of target materials. Such technicalities cost a significant amount of human time and expose practitioners to mistakes. Here, we introduce a fully automated package for DFT calculations, automated *ab initio* modeling of materials property package (AMP<sup>2</sup>), which aims to produce key DFT properties of crystalline materials with essentially no user intervention except for initial structural information. This is achieved through algorithms that automatically determine various technical parameters and make self-decisions during computational workflow. As results, AMP<sup>2</sup> is material-agnostic and provides a highly accurate band structure, band gap, effective mass, density of states and dielectric constant for the given material. Notably, the package finds the antiferromagnetic ground state by applying a genetic algorithm to effective Ising models. AMP<sup>2</sup> also addresses band-gap underestimation in semilocal functionals with help of a hybrid functional, thereby producing a more accurate band gap, even if the material turns out to be metallic within the semilocal functional. We believe that the present package will significantly expand usage of DFT calculations by making them more accessible.

### Program summary

Program Title: AMP<sup>2</sup>

CPC Library link to program files: <http://dx.doi.org/10.17632/5rdw9jv5vp.1>

Licensing provisions: GPLv3

Programming language: Python

Nature of problem: Conducting *ab initio* calculation under a fully automated protocol to obtain crystalline properties

Solution method: Construct workflow to estimate material properties using the density functional theory. Ising model and genetic algorithm are used for identifying magnetic spin ordering.

© 2020 Published by Elsevier B.V.

## 1. Introduction

Owing to the high accuracy and *ab-initio* nature, theoretical analysis based on the density functional theory (DFT) [1] is now established as a principal pillar in material science, explaining experimental data in atomistic detail. The predictive capability

of DFT also enables screening a large pool of materials, thereby suggesting promising candidates in various applications [2–13]. The popularity of DFT calculations is largely indebted to sophisticated programs such as VASP [14], Quantum-ESPRESSO [15], SIESTA [16], ABINIT [17], CASTEP [18], etc. These programs become reliable through enduring tests by a vast number of research groups from diverse fields.

While many DFT programs constantly improve to be more user-friendly, the *ab initio* calculation still requires in-depth knowledge on underlying numerical algorithms as well as fundamental theories on electronic structures. This is the case even for calculations on simple crystals, which forms a base for more advanced studies on defects and surfaces. For instance, several

<sup>☆</sup> The review of this paper was arranged by Prof. Stephan Fritzsche.

<sup>☆☆</sup> This paper and its associated computer program are available via the Computer Physics Communication homepage on ScienceDirect (<http://www.sciencedirect.com/science/journal/00104655>).

\* Corresponding author.

E-mail address: [hansw@snu.ac.kr](mailto:hansw@snu.ac.kr) (S. Han).

computational parameters should be chosen by considering whether the material is metallic or insulating. Furthermore, DFT calculations undergo several intermediate steps before obtaining a physical quantity, demanding careful management of input or output files. These technical hurdles stretch the learning period that practitioners should spend on accumulating enough experience to be able to produce reliable data.

Recently, several automation utilities for DFT calculations have been developed with an aim to accelerate data production by providing algorithms for calculating target properties. (For instance, AiiDA [19], atomate [20], pymatgen [21], and aflow [22].) While these automation packages help lower the learning curve as well as reduce human mistakes in using DFT programs, current implementations still assume that users are familiar with technical details, which can be a significant barrier to less experienced users.

In this paper, we introduce ‘automated *ab initio* modeling of materials property package’ (AMP<sup>2</sup>), a full-fledged automation script running a DFT program, specifically VASP. Starting from information only on the crystalline structure, AMP<sup>2</sup> calculates key properties of crystalline materials using algorithms that automatically determine various technical parameters and make self-decisions during the computational workflow. The main philosophy of the package is to minimize human interference in *ab initio* calculations, still producing highly accurate results. This is enabled by sophisticated algorithms that incorporate well-established know-hows, thereby fully automating the computational procedure.

AMP<sup>2</sup> is material-agnostic and requires only the structural information from the user, and delivers various crystalline properties such as band gap ( $E_g$ ), band structure, density of states (DOS), effective masses, and dielectric constants (optical or static). Notably, the package finds the antiferromagnetic ground state and overcomes band-gap underestimation in conventional DFT calculations with the help of a hybrid functional. The crystalline properties obtained by AMP<sup>2</sup> will be useful in advanced materials modeling on defects and surfaces.

The following sections are organized as follows: in Section 2, we overview the main workflow of AMP<sup>2</sup>. Section 3 explains key algorithms in more detail. In Section 4, we show how to install and use AMP<sup>2</sup> with discussions on test results and limitations. Finally, we summarize and conclude in Section 5.

## 2. Overall workflow

Fig. 1 shows the main workflow of AMP<sup>2</sup>. The DFT calculations in the package are carried out by VASP [14], and users should have a valid version of VASP installed in the local computer. Since most functions required by AMP<sup>2</sup> are also available in other DFT codes, the present automation framework would be also compatible with them via proper modifications. (For instance, we were able to make a version of AMP<sup>2</sup> to be compatible with Quantum-Espresso just by developing interfacing modules between Quantum-Espresso and common libraries.) First, the code receives a structure file that contains information on atomic positions and lattice vectors. The valid structure formats are Crystallographic Information File (CIF) or VASP input files. Subsequent workflows are split into two parts; core setting and computation of physical properties. In the core setting, the package first carries out convergence tests to determine basic computational parameters such as  $\mathbf{k}$  points and cutoff energy ( $E_{\text{cut}}$ ), then subsequently obtain the theoretical equilibrium structure. In particular, the package identifies zero-temperature, collinear magnetic ordering for materials possessing local magnetic moments. While there is no restriction on input structures, it is implicitly assumed that the given structure is reasonably

close to the theoretical equilibrium. This is because convergence tests are carried out only on the initial structure. About the exchange–correlation functional, the default setting uses Perdew–Burke–Ernzerhof (PBE) functional [23] as a generalized-gradient approximation (GGA) but the local-density approximation (LDA) [24] and hybrid-functional method [25] can be used as well. By default, the spin–orbit-coupling (SOC) is not included but users can activate the SOC option when necessary. (The effects of SOC can be significant for elements heavier than Tl.) For electron-ion interactions, basic types of projector-augmented wave (PAW) pseudopotentials are selected but users can choose other types if necessary. Detailed guidelines to modify the computational setting are discussed in the manual [26].

Computational conditions and equilibrium structure determined in the core setting transfer to ensuing calculations on crystalline properties. The electronic information such as the band structure and DOS is obtained first. The underestimation of  $E_g$  is addressed by the one-point hybrid functional method [4]. (See below for details.) For materials with a relatively small gap ( $\lesssim 1$  eV) such as Ge, the PBE functional often concludes a metallic nature incorrectly. To overcome such problems that are typical in semilocal functionals, AMP<sup>2</sup> examines DOS near the Fermi level ( $E_F$ ) for metallic materials within PBE and inspect a possibility of gap opening. By carrying out one-point hybrid-functional calculations on selected  $\mathbf{k}$  points, the package determines the metallic or semiconducting nature more reliably. For materials with a finite  $E_g$ , the package continues to evaluate dielectric constants and average effective masses of electrons or holes. By default, all the properties are calculated but users can choose target quantities selectively.

For transition-metal elements with localized 3d electrons (V ~ Zn), electronic correlations from the on-site Coulomb repulsion can significantly affect the band structure. This is efficiently described by PBE + Hubbard  $U$  method [27]. We adopt  $U$  values from Ref. [28]. However, for metals where the on-site Coulomb interaction is screened by free carriers, it is more appropriate to use the bare PBE functional ( $U = 0$ ). In order to take this into account, AMP<sup>2</sup> applies the PBE+ $U$  scheme only when the compound simultaneously includes 3d transition-metal element (V~Zn) and non-metal (H, He, C, N, O, F, Ne, P, S, Cl, Ar, Se, Br, Kr, I and Xe) or semi-metal (B, Si, Ge, As, Sb, and Te) elements. Furthermore, if materials are found to be metallic from the band structure and small-gap correction (see above), then the calculation restarts from the beginning without on-site  $U$  terms. (Nevertheless, the results from PBE+ $U$  are stored because they can be used later in evaluating the formation energies of related compounds that need  $U$  contributions.) For half-metals such as CrO<sub>2</sub>, Fe<sub>3</sub>O<sub>4</sub>, Co<sub>2</sub>MnSi, and Li<sub>2</sub>NiSiO<sub>4</sub>, it is known that correct electronic structures are obtained by including on-site Coulomb interactions [29,30]. Therefore, AMP<sup>2</sup> applies PBE+ $U$  methods to this class of materials in spite of metallicity.

## 3. Key algorithms in AMP<sup>2</sup>

### 3.1. Convergence test for $k$ points and cutoff energy

In solving Kohn–Sham equations within the momentum-space formalism [31] using a plane-wave basis, one should introduce approximations on the  $\mathbf{k}$ -point mesh in the Brillouin-zone (BZ) integration and the maximum kinetic energy ( $E_{\text{cut}}$ ) for the plane-wave expansion. These parameters are usually selected by inspecting convergence of various physical properties. For the core setting, AMP<sup>2</sup> examines the total energy per atom and stress tensor, and determines parameters that are reasonable for most calculations. We note that computational costs for the convergence test are negligible compared to other steps. In addition,

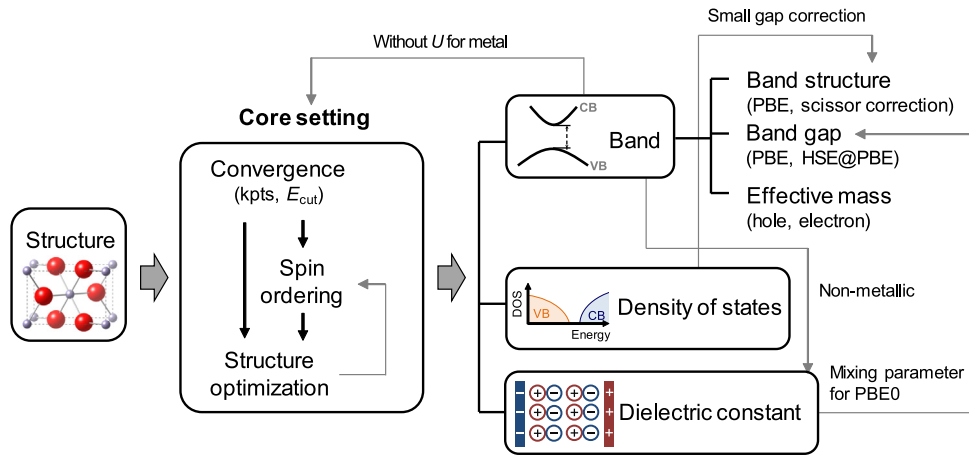


Fig. 1. The main workflow of AMP<sup>2</sup>. The light arrows indicate the flow of information between different types of calculations.

different cutoff energies are often required between similar materials (for example, AlCaF<sub>5</sub> and AlCa<sub>2</sub>F<sub>7</sub>). Therefore, it is safe and reasonable to perform convergence tests independently for each material.

The  $\mathbf{k}$  points are selected on a regular mesh within the first BZ [32]. The dimension of mesh grids ( $N_1, N_2, N_3$ ) and fractional  $\mathbf{k}$  points along each direction are determined by an integer parameter  $n_k$  as follows:

$$N_i = \text{ceil}[b_i/(b_{\max}/n_k)] \quad (i = 1, 2, 3), \quad (1)$$

$$k_i^m = \frac{2m - N_i - 1}{2N_i} \quad (m = 1, 2, \dots, N_i), \quad (2)$$

where  $i$  is the axis index,  $b_i$  is the length of  $i$ th reciprocal lattice,  $b_{\max}$  is the maximum among  $\{b_i\}$ . According to Eq. (1),  $N_i$  for the longest reciprocal lattice is  $n_k$  while other  $N_i$ 's are scaled by  $b_i/b_{\max}$ . By ratcheting up  $n_k$ , more dense meshes are selected. To note, if  $N_i$  is an odd number, 0 is selected in  $k_i^m$  along the  $i$ th axis while it lies on the middle of the grid for an even  $N_i$ . For the BZ integration, the package uses a Gaussian smearing with the width of 0.05 eV, which gives reasonable results in most cases.

The  $\mathbf{k}$ -point meshes in Eqs. (1) and (2) work well for most materials. However, some symmetry groups need caution; in the case of hexagonal and rhombohedral structures, AMP<sup>2</sup> shifts the  $\mathbf{k}$ -point mesh in Eq. (2) such that the mesh always includes the  $\Gamma$  point. This is to ensure that symmetrically equivalent  $\mathbf{k}$  points lie on the mesh. For the same reason, an overriding rule applies to the body-centered tetragonal and body-centered orthorhombic cells. In these unit cells, axis lengths of the primitive cell are identical but the reciprocal lattice vectors differ in length. However,  $N_i$ 's are equalized to make the  $\mathbf{k}$ -point mesh to be self-inclusive under the symmetry operations.

The PAW pseudopotentials include  $E_{\text{cut}}$  recommended for each element, and VASP uses the maximum  $E_{\text{cut}}$  among constituent elements by default. Starting from an  $E_{\text{cut}}$  that is 50 eV below the VASP default, AMP<sup>2</sup> increases  $E_{\text{cut}}$  by 50 eV. As shown in Fig. 2, the package determines a converged  $\mathbf{k}$ -point grid and  $E_{\text{cut}}$  when the total energy and stress-tensor components agree with next two values within 10 meV/atom and 10 kbar (the next values also should agree with each other). The present convergence test usually results in  $E_{\text{cut}}$  higher than the VASP default by 50–100 eV. The energy difference, which is a more relevant quantity, converges faster than the total energy. Therefore, under the convergence criterion of 10 meV/atom, we expect that energy differences converge within a few meV/atom, which would be sufficient in most cases. On the other hand, the convergence condition of 10 kbar on stress tensors considered that typical

Young's moduli of inorganic materials are over 100 GPa (1 Mbar), which implies that errors in lattice parameters are less than 1%. However, soft materials call for cautions (see Section 4.4).

We do not consider the convergence in atomic forces because they are strictly zero in the crystals with high symmetry. Nevertheless, we find that the convergence of energy and stress also ensures that of atomic force. For example, Fig. 2 shows that atomic forces converge within 0.02 eV/Å. We also note that the converged  $\mathbf{k}$ -point grids are rather insensitive to  $E_{\text{cut}}$  as confirmed in Fig. S1 and S2, implying that the convergence behaviors with respect to the  $\mathbf{k}$ -point grid and  $E_{\text{cut}}$  are decoupled. Therefore, the  $\mathbf{k}$ -point test is done first with a default  $E_{\text{cut}}$  followed by the  $E_{\text{cut}}$  test at the converged  $\mathbf{k}$ -point grid.

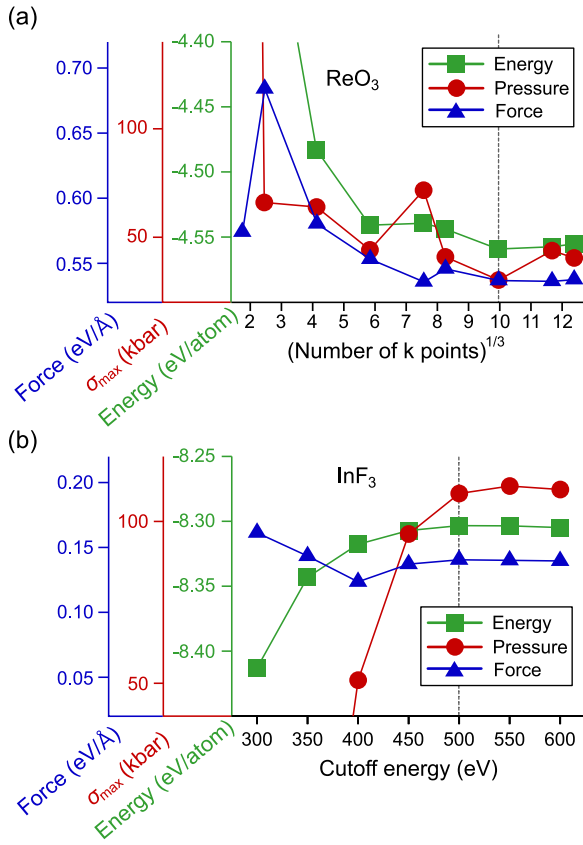
### 3.2. Optimizing structures

To obtain a structure that is theoretically optimized, the structural relaxation is conducted until magnitudes of atomic forces and every component of stress tensors are lower than 0.02 eV/Å and 10 kbar, respectively. During the relaxation, the initial symmetry is maintained. For each ionic step, the self-consistent iteration is carried out until the total energy converges to within  $10^{-6}$  eV. Following the default setting in VASP, the package employs the Block Davidson method for every procedure. If the total energy does not converge in spin-polarized calculations, the package modulates parameters for mixing input and output charge densities according to a recipe suggested in the VASP manual [33].

Occasionally, the lattice parameters change significantly during optimization. This effectively alters  $E_{\text{cut}}$  because the dimension of real-space grids is fixed. To maintain the computational accuracy, the package iterates the structural relaxation from the previously optimized structure until lattice vectors change lengths and angles within 0.2% and 0.1°, respectively, after relaxation.

### 3.3. Magnetism

By default, AMP<sup>2</sup> assumes spin-unpolarized calculations. However, for elements possessing localized or unpaired electrons, the spin-polarization would be significant. To address this, the package automatically turns on ferromagnetic spin-polarization if input materials satisfy any of the two following conditions; (i) Materials include metal elements whose valence electrons partially occupy d orbitals (Group 4 ~ Group 12) or f orbitals (lanthanides and actinides). (ii) The total number of electrons in the unit cell is odd. The electronic iteration requires starting

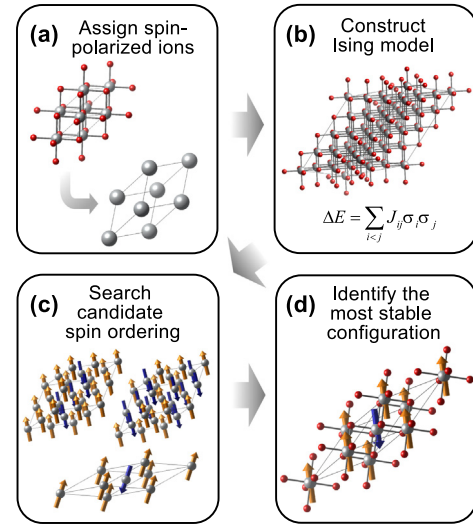


**Fig. 2.** Convergence test of (a)  $\mathbf{k}$ -points for ReO<sub>3</sub> and (b) cutoff energy for InF<sub>3</sub>. The vertical dashed lines indicate the converged  $\mathbf{k}$ -points (or cutoff energy).  $\sigma_{\max}$  is the maximum absolute component in the stress tensor. For comparison, the maximum force component is also plotted.

magnetic moments of atoms to construct an initial spin-polarized charge density. In the case of (i), the package sets initial magnetic moments considering the number of valence electrons and charge states (if available) in CIF files. This usually helps facilitate the electronic convergence. For example, the initial magnetic moment of Ni<sup>2+</sup> in NiO is set to be  $(0.5 + 2) \mu_B$  (2 is the number of unpaired electrons in the high-spin configuration). If the information on the charge state is not available, the default value of  $3 \mu_B$  is used. In the case of (ii), the package uses a default setting in VASP ( $1 \mu_B$  for every element).

The spin-polarized calculations are performed for convergence tests and structural relaxations. Then, if the local magnetic moment of every atom is less than  $0.2 \mu_B$ , AMP<sup>2</sup> concludes that the material is non-magnetic and carry out additional relaxations with the spin-unpolarized setting. In estimating local magnetic moments, we employ a projection scheme utilizing PAW projectors (LORBIT = 11). The threshold value of  $0.2 \mu_B$  is chosen on the following grounds; for 100 materials in which every local moment is less than  $0.2 \mu_B$  under a spin-polarized initial setup, we find that energy differences between spin-unpolarized and spin-polarized calculations are less than 3 meV/atom with almost identical DOS, indicating a non-magnetic nature.

Notably, AMP<sup>2</sup> provides an option to search a stable anti-ferromagnetic ordering. This is particularly useful for materials including 3d transition metal atoms (Ti~Cu). The package finds the ground-state collinear magnetic spin ordering using the Ising model, following an algorithm suggested in Ref. [34]. Fig. 3 illustrates the schematic procedure. First, ions with local magnetic moments are assigned [Fig. 3(a)] based on the starting ferromagnetic calculations. Next, a supercell to evaluate Ising coefficients



**Fig. 3.** Schematic diagram for identifying the ground state spin ordering with the example of NiO.

is constructed and the package estimates exchange integrals from total energies of various spin configurations [Fig. 3(b)]. (Here, magnitudes of initial magnetic moments are set to those from the ferromagnetic calculation.) Once every Ising coefficient is determined, supercells within the Ising model are constructed by expanding the unit cell in various ways up to a  $2 \times 2 \times 2$  repetition. To search for low-energy spin ordering, AMP<sup>2</sup> considers every possible spin configuration if the number of spin-polarizable ions is less than 9. Otherwise, a genetic algorithm is used. The candidate orderings with low energies are then selected [Fig. 3(c)] and their primitive cells are identified by utilizing the spglib library in Phonopy [35]. Finally, by comparing DFT total energies among the candidate structures, AMP<sup>2</sup> finds the most stable ordering [Fig. 3(d)]. In doing so,  $\mathbf{k}$  points are resampled for each structure according to Eqs. (1) and (2) adopting the converged  $b_{\max}/n_{\mathbf{k}}$  for the ferromagnetic cell. For computational consistency, AMP<sup>2</sup> obtains a converged  $\mathbf{k}$ -point grid freshly for the final antiferromagnetic unit cell. In Ref. [34], the present algorithm successfully identified ground-state spin ordering of various transition metal oxides such as NiO, Mn<sub>3</sub>O<sub>4</sub> and CoCr<sub>2</sub>O<sub>4</sub>. We also confirm that the method works well for antiferromagnetic metals such as MnPt, MnS, FeCl<sub>2</sub>, and FeF<sub>2</sub>. To note, the additional computational cost for identifying the anti-ferromagnetic structure is not significant compared to the time for evaluating the whole properties.

### 3.4. Band structure and density of states

For the band structure, AMP<sup>2</sup> computes Kohn–Sham eigenvalues along the lines connecting every combination of high-symmetry  $\mathbf{k}$  points in Ref. [22] with a line density of  $0.04\pi \text{ \AA}^{-1}$  by carrying out non-self-consistent calculations. We name the selected  $\mathbf{k}$ -point set as  $\mathbb{K}$ . In graphic presentation of band structures, those paths explicitly chosen in Ref. [22] are displayed. The Fermi level ( $E_F$ ) is obtained during self-consistent calculations ( $E_F^{\text{SCF}}$ ). In the case of metallic systems,  $E_F$  is simply equal to  $E_F^{\text{SCF}}$ . For semiconductors or insulators,  $E_F$  is set to the maximum eigenvalue within valence bands (enumerated for  $\mathbb{K}$ ) because sparse sampling in the self-consistent calculation could underestimate  $E_F$  for dispersive valence bands. (The next subsection discusses how to judge metallicity.) In obtaining DOS, the package doubles the mesh dimension of  $\mathbf{k}$  points compared to that used in the structural optimization, which provides almost converged



DOS shapes (see Fig. S3 for comparing DOS between doubling and quadrupling  $\mathbf{k}$  points). The package calculates DOS with the tetrahedron method with the Blöchl correction [36]. AMP<sup>2</sup> does not include spin-orbit coupling (SOC) by default but users can turn it on for the band structure and DOS by specifying elements. In this case, the self-consistent calculation is also performed with SOC included.

### 3.5. Band gap

In order to determine metallicity or estimate  $E_g$ , AMP<sup>2</sup> utilizes the band structure obtained in the above. The package first sorts all the eigenvalues evaluated at  $\mathbf{k} \in \mathbb{K}$  in the increasing order and occupies them up to  $N_{el} \cdot N_k \cdot N_{spin}$  where  $N_{el}$  is the number of electrons in the unit cell,  $N_k$  is  $\#\mathbb{K}$ , and  $N_{spin}$  is the spin degree of freedom (1 or 2). Then,  $E_F^{band}$  is defined as an average between the highest energy among occupied states and the lowest energy among unoccupied states. If  $E_F^{band}$  lies between the maximum and minimum energies of any one band (i.e., the same band index), the material is obviously metallic because  $E_F$  should cross the band. Otherwise, the valence-band maximum (VBM) and conduction-band minimum (CBM) are set to the energy level closest to  $E_F^{band}$  among the occupied and empty bands, respectively. If  $E_g = E_{CBM} - E_{VBM} < 0.01$  eV, the material is regarded as a metal, otherwise it is an insulator or semiconductor with the band gap of  $E_g$ . (The finite upper bound of 0.01 eV considers that eigenvalues are calculated at discrete points, not along continuous lines.)

It is well known that the PBE functional underestimates the band gap by 40%–60% [37]. The hybrid functional method, notably Heyd–Scuseria–Ernzerhof functional (HSE06; simply HSE in this work) [25] significantly alleviates the discrepancy by partially incorporating exact exchange energies, thereby mitigating self-interaction errors in the conventional DFT. However, the method is computationally too expensive to apply to the whole procedure. Thus, AMP<sup>2</sup> estimates HSE eigenvalues at  $\mathbf{k}$  points corresponding to VBM and CBM from the PBE band structure while self-consistent calculations are carried out within HSE based on the relaxed structure in the PBE calculations. We call this approach one-shot hybrid calculations (HSE@PBE). This assumes that band edges from HSE and PBE lie on the same  $\mathbf{k}$  points. In Ref. [4], it was shown that this approach yields band gaps in many materials similar to those from the full HSE calculations. The mixing parameter for the exact exchange energy is fixed to 0.25 by default but users can optionally change the fraction to  $\epsilon_0^{-1}$  within PBE0 [38], which improves the correlation with experimental data (see Fig. S4) [39]. In the HSE@PBE calculations, the package employs a  $\mathbf{k}$ -point mesh that converges only the total energy instead of both energy and stress tensor, which shrinks mesh dimensions and hence computational costs of the hybrid functional. We note that this does not affect the band gap because the structure optimization is not performed in the HSE@PBE scheme.

### 3.6. Identification of small-gap materials

Due to the band-gap underestimation, small-gap semiconductors such as Ge and InAs are calculated to be metallic in PBE. To distinguish small-gap materials from true metals, AMP<sup>2</sup> employs an indicator,  $D_F/D_{VB}$ , that is defined as follows:

$$D_F/D_{VB} = \frac{\int_{-0.3}^{0.3} D(E) dE}{0.6} \bigg/ \frac{\int_{-3}^{-1} D(E) dE}{2}, \quad (3)$$

where  $D(E)$  is DOS at the energy of  $E$  when  $E_F = 0$ . All the numbers in Eq. (3) are in eV. Here, AMP<sup>2</sup> uses DOS obtained with a Gaussian smearing to avoid sharp features in the tetrahedron method. As illustrated schematically in Fig. 4(a), suppressed DOS

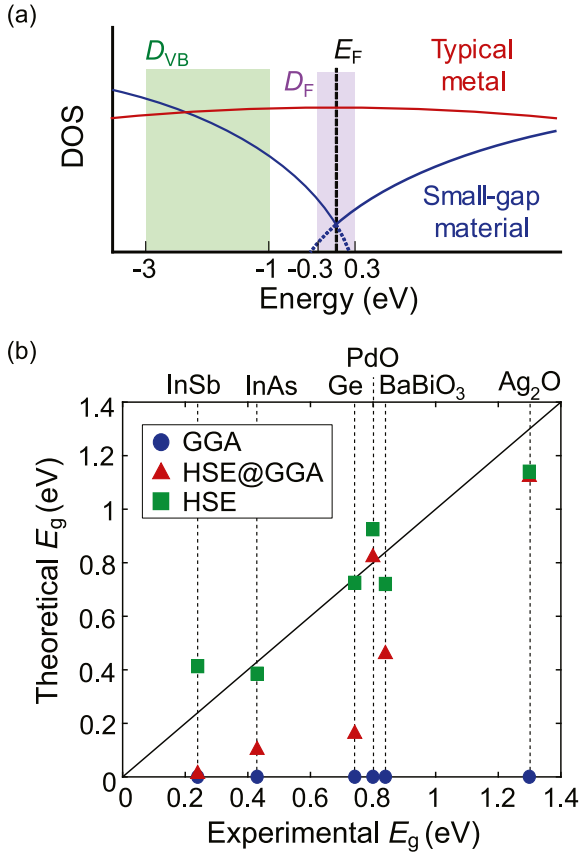
near  $E_F$  with small  $D_F/D_{VB}$  indicate a possibility that the valence and conduction bands overlap due to the band-gap underestimation. The integration range and threshold value for  $D_F/D_{VB}$  were carefully tuned over 60 binary oxides that are metallic within PBE. By choosing an integration boundary of  $[-3, -1]$  for  $D_{VB}$ , we can set a threshold of 0.3 for  $D_F/D_{VB}$  indicators above which all materials are metal and below which 80% is found to be small-gap oxides in hybrid calculations. Additionally, we confirm that the method can successfully apply to other 20 ternary oxides and 20 binary compounds with  $D_F/D_{VB}$  around 0.3.

For materials with  $D_F/D_{VB} < 0.3$ , AMP<sup>2</sup> first identifies local extrema in the band structure within  $\pm 0.5$  eV from  $E_F$ . Assuming that the  $\mathbf{k}$  points for local extrema are the same between PBE and HSE, the package carries out HSE@PBE calculations on  $\mathbf{k}$  points corresponding to local extrema. Afterwards, the metallicity or band gap is determined following the same method in the previous subsection. If the material still turns out to be metallic, only PBE results are presented. (Since spin-unpolarized or anti-ferromagnetic materials with odd number of electrons are always metallic, they do not undergo the present step.) Fig. 4(b) shows results for selected materials. It is seen that the materials that are incorrectly classified as metals within PBE exhibit finite gaps under the present approach. However, the gap opening is not sufficient for several materials. We note that this is not caused by limitations in the present method. Rather, the error originates from the lattice parameters overestimated by PBE. When full HSE calculations are performed, the agreements with the experiment are much better [see squares in Fig. 4(b)]. For options enabling the lattice relaxation within HSE, we refer to Section 3.3.1 of the manual.

AMP<sup>2</sup> can falsely report an insulating gap even though the given material is metallic within a full HSE calculation. This can occur when the band does not rigidly shift between PBE and HSE such that band extrema locate at different  $\mathbf{k}$  points between the two functionals. However, we could not find such errors for 20 test materials in which gap opens under the present correction scheme. On the other hand, it is still possible that the HSE functional itself misjudges a metal as an insulator while PBE may correctly predict the metallicity. Although we are not aware of any such example, one cannot exclude its possibility.

### 3.7. Gap-corrected band structure

Additionally, AMP<sup>2</sup> provides a gap-corrected band structure by applying a scissor correction that shifts up conduction bands in PBE rigidly to match with  $E_g$  from HSE@PBE. In materials with a finite  $E_g$  in PBE, it is straightforward to distinguish valence and conduction bands and the scissor correction can be easily applied. In small-gap materials that are metallic within PBE, however, it is tricky to determine the band type (valence or conduction) particularly near  $E_F$ . Following a method used in Quantum-ESPRESSO [15], the package identifies the band type by reordering band indexes based on overlap integrals between wavefunctions with eigenvalues within  $\pm 3$  eV from  $E_F$ . Then, a band is classified as the conduction band if its energies are higher than  $E_F + 0.5$  eV at most  $\mathbf{k}$  points. [See Fig. 5(a) for the example of PdO.] Then, the package shifts the conduction bands to match the band gap with the  $E_g^{HSE@PBE}$ . As an example, the gap-corrected band structure is shown in Fig. 5(b) for PdO, which is in good agreement with the full HSE band structure. From various tests, we find that the above algorithm works in most materials except for some cases where orbital characters between conduction and valence bands are highly mixed. (One example for CaPd<sub>3</sub>O<sub>4</sub> is provided in the Supplementary Material.) In this case, the package does not provide the modulated band structure.



**Fig. 4.** (a) Schematic density of states (DOS) for typical metal and small-gap material that is metallic within PBE. Shade regions for  $D_{VB}$  and  $D_F$  indicate the integration ranges in Eq. (3). (b) Comparison of theoretical and experimental band gaps for small-gap materials [40–43].

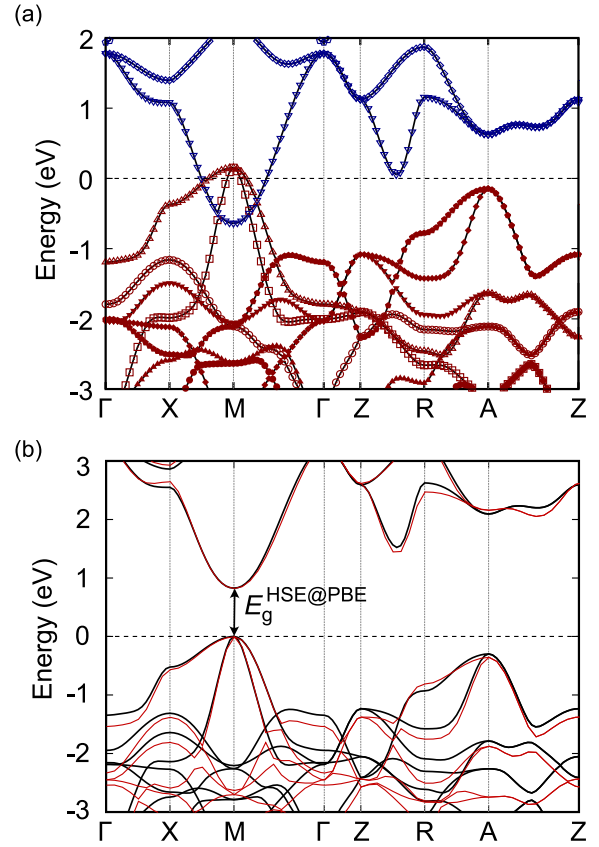
### 3.8. Effective mass

For insulators and semiconductors, AMP<sup>2</sup> computes averaged effective-mass tensors ( $\langle m_{ij}^* \rangle$ ) within the semiclassical transport theory using the following equation:

$$\langle m_{ij}^* \rangle^{-1} = \frac{1}{\hbar^2} \frac{\sum_{n,\sigma} \int d\mathbf{k} \frac{\partial^2 \varepsilon_{n\mathbf{k}\sigma}}{\partial k_i \partial k_j} f_{n\mathbf{k}\sigma}}{\sum_{n,\sigma} \int d\mathbf{k} f_{n\mathbf{k}\sigma}}, \quad (4)$$

where  $\hbar$  is the Planck constant divided by  $2\pi$ ,  $n$  the index for valence (conduction) bands for electrons (holes),  $\mathbf{k}$  the reciprocal wave vector,  $\sigma$  the spin index,  $\varepsilon_{n\mathbf{k}\sigma}$  the corresponding Kohn–Sham eigenvalue, and  $f_{n\mathbf{k}\sigma}$  is the Fermi–Dirac distribution function ( $f_{FD}$ ) at 300 K for electron carriers or  $1 - f_{FD}$  for hole carriers. In Eq. (3),  $i$  and  $j$  denote Cartesian indices ( $x$ ,  $y$ , and  $z$ ). (The temperature can be changed as instructed in the manual.) The Fermi level in  $f_{FD}$  is set to CBM (VBM) in evaluating the electron (hole) effective mass.

In order to reduce computational costs without compromising accuracy, we restrict the  $\mathbf{k}$  domain in Eq. (3) by considering that  $f_{n\mathbf{k}\sigma}$  rapidly decays to zero for  $\varepsilon_{n\mathbf{k}\sigma}$  away from the Fermi level. This scheme contrasts with other programs, for example, Boltztrap [44] that utilizes  $\mathbf{k}$ -point meshes over the whole BZ. The relevant regions are selected by the following three steps (it is assumed that hole effective masses are calculated): first, the package locates local extrema along the  $\mathbf{k}$ -point lines explored in the band structure. If the energy of the top valence band at a certain  $\mathbf{k}$  is the maximum within a cutoff radius of  $0.04\pi \text{ \AA}^{-1}$  and  $f_{n\mathbf{k}\sigma} > 0.01$ , the  $\mathbf{k}$  point is chosen to be a local maximum. Then,

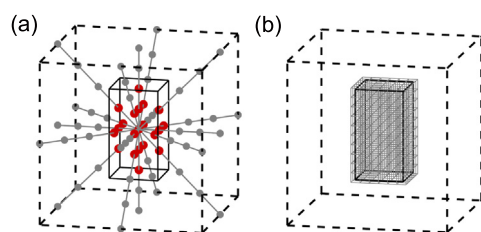


**Fig. 5.** (a) Band structure of PdO calculated within PBE. The energy eigenvalues are connected with the same symbol considering the overlap integral. The blue and red lines indicate the conduction and valence bands, respectively. (b) The corrected band structure (black lines) that is obtained by rigidly shifting conduction and valence bands in (a) and fitting the band gap to  $E_g^{\text{HSE@PBE}}$ . The full hybrid-functional band structure is shown in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the package examines eigenvalues of the corresponding band along  $\mathbf{k}$  lines in directions of  $\langle 100 \rangle$  (face center) and  $\langle 110 \rangle$  (edge center) that pass through the extremum point [total 9 lines; see Fig. 6(a)]. Then, the package determines the smallest orthorhombic cell that encloses all points with  $f_{n\mathbf{k}\sigma} > 0.01$ . If several extremum points are close to each other, the algorithm chooses one region that includes all of them. After constructing a uniform mesh with the spacing of  $0.04\pi \text{ \AA}^{-1}$ , the package carries out differentiation and integration in Eq. (3) based on centered difference approximations. The integration in the numerator of Eq. (3) is carried out around extremum points in the irreducible BZ, and the package expands the result over the 1st BZ through symmetry operations. In Fig. S6, we compare hole effective masses from AMP<sup>2</sup> and those from obtained using Boltztrap program [7]. While AMP<sup>2</sup> uses typically several hundreds of  $\mathbf{k}$  points, far less than 8000 points in Ref. [7], the agreements are excellent.

### 3.9. Dielectric constant

There are two types of dielectric tensors depending on whether ions are clamped or allowed to relax under an external field; the optical dielectric constant ( $\varepsilon^\infty$ ) and the static dielectric constant ( $\varepsilon^0$ ), respectively. VASP produces both quantities using a linear-response method based on the density functional perturbation theory (DFPT) [45] and AMP<sup>2</sup> simply quote the results. Since dielectric constants are sensitive to the  $\mathbf{k}$ -point



**Fig. 6.** The schematic diagram to build  $k$ -space grids for calculating the effective mass. (a) The restricted mesh grid space (black box with the solid line). The computed points (gray and red dots) are determined along the specific lines in gray. The red dots indicate points at which the hole occupation is bigger than 0.01. (b) The mesh grid for finite difference method.

sampling, the package doubles the  $k$ -point density along each direction. Also, the self-consistency threshold in electronic steps is tightened to  $10^{-8}$  eV. Since the dielectric calculation is only meaningful for non-metallic compounds, the calculation is not performed if the material is confirmed to be metallic within PBE. We note that the PBE functional usually overestimates the dielectric constant, in particular for high dielectric constant (high- $k$ ) materials [4]. Therefore, we recommend using LDA for high- $k$  materials, which can be done by simply switching the functional from the beginning. (See the manual for detailed instructions.) Occasionally, imaginary phonon modes are observed at the zone center for various reasons, for example, structural instability. In these cases, the package does not provide  $\epsilon^0$ .

## 4. Usage

### 4.1. Installation and execution

AMP<sup>2</sup> can be downloaded from <https://github.com/MDIL-SNU/AMP2>. The code supports both Python2.7 and Python3. In addition, it requires several python modules such as 'math' and 'numpy' for numerical calculations as well as 'scipy' for various physical constants. AMP<sup>2</sup> utilizes spglib [35] to identify symmetry operators and primitive cell, and gnuplot for drawing figures, so these programs should be also pre-installed. The DFT calculation in AMP<sup>2</sup> is performed by the VASP package (AMP<sup>2</sup> has been thoroughly tested with version 5). Users start AMP<sup>2</sup> by executing `run.sh` that specifies system setting such as parallel environment and job scheduler.

### 4.2. Input files and parameters

As input files, AMP<sup>2</sup> requires a file for the crystal structure and `config.yaml` that informs computational options. The valid structure formats are VASP input files (POSCAR) and cif format. For the cif format, AMP<sup>2</sup> requires a standard style [46] that includes symmetry operators (`_space_group_symop_[]` or `_symmetry_equiv_[]`), atomic labels (`_atom_site_label`), occupancy (`_atom_site_occupancy`) and fractional positions (`_atom_site_fract_[]`). The structure file should be labeled as `name.cif` or `POSCAR_name` where `name` is a user-defined string and will be used in naming directories (see below). On the other hand, `config.yaml` should provide computational environment for AMP<sup>2</sup> such as paths for binary executables and directories containing the structure file and source codes of AMP<sup>2</sup>.

With the minimal input from users explained above, AMP<sup>2</sup> calculates all the properties in the previous section. Users can selectively turn on/off the computation of each property within `config.yaml`. However, the effective mass and band gap with HSE@PBE can be performed after the band calculation is finished

(see Fig. 1). Users can also turn on/off SOC using `SOC_target` tag in `config.yaml`. If there exist multiple structure files in the directory, AMP<sup>2</sup> carries out calculations on all structure files and store results under different directories. This feature will be useful for high-throughput computing.

### 4.3. Directory trees for output

The computational results are stored under repository directories with names reflecting that of the structure file (see above). There are various sub-directories that correspond to target properties. The list of these directories and key files are summarized in Table 1. INPUT0 is the main directory wherein all the input files for DFT calculations are stored. If the antiferromagnetic ordering is the most stable spin configuration, input files are updated to reflect the spin ordering after magnetic-ordering step and previous input files are saved under `INPUT0_old`. Convergence tests for  $k$  points and cutoff energy are stored in sub-directories named `kptest` and `cutoff`, respectively. The results for convergence test are written in `kpoint.log` and `cutoff.log` and users can plot the convergence behavior like in Fig. 2 using the input file for gnuplot (`conv_plot.in`). On the other hand, `magnetic_ordering` is the path where the procedure to find ground state magnetic ordering is saved. The configuration with the ground state spin ordering is located in `INPUT0/POSCAR_rlx_GGA`.

In each directory, the package records checkpoints for the calculation within the file `amp2.log`. (`amp2.log` in the main directory notes the whole procedure.) Users can select the exchange–correlation functional such as PBE or LDA and the directory name depends on the potential such as `relax_GGA` and `relax_LDA`. The optimized structure is stored in `CONTCAR` file in `relax_GGA` (or `relax_LDA`) directory. Under `band_GGA`, the band gap is written in `Band_gap.log` and the figure for band structure can be drawn using `band.in`. Additionally, the band gap within HSE@PBE is written in `Band_gap.log` in `hybrid_GGA` directory. The density of states can be plotted using `dos_GGA/Pdos_dat/dos.in`. The dielectric tensor is recorded in `dielectric.log` in `dielectric_GGA` and the effective mass tensor for hole and electron are written in `effective_mass.log` in `effm_GGA/hole` and `effm_GGA/electron`, respectively. All the main results including log and image files are compiled in `Results`.

### 4.4. Tests and limitations

We started to develop AMP<sup>2</sup> in 2012 and have employed AMP<sup>2</sup> in automated high-throughput screenings on high- $k$  materials [4,5] and p-type oxides [8,9]. During the extensive applications, we constantly updated AMP<sup>2</sup> whenever the program failed (bad convergence or inaccurate results). In total, AMP<sup>2</sup> was tested on  $\sim 7000$  materials that encompass various types of materials including  $\sim 3000$  oxides. Computational results on these materials are open on a website (<https://www.snumat.com>) [47]. Most functions in AMP<sup>2</sup> have been thoroughly tested except for two newest features: identifying small-gap materials and antiferromagnetic ordering. Thus, we freshly tested the up-to-date version of AMP<sup>2</sup> against two groups of materials: the first group consists of well-known, archetypal materials: Al, Cu, Fe, SnO, ZnO, Ge, InAs, Si, AlN, ZnS, SiC, GaP, GaN, AlP, Si<sub>3</sub>N<sub>4</sub>, SiO<sub>2</sub>, TiO<sub>2</sub>, CrO<sub>2</sub>, NiO and Cr<sub>2</sub>O<sub>3</sub>, covering a wide range of band gaps, dielectric constants, effective masses, and magnetic ordering. We find that AMP<sup>2</sup> successfully finishes every calculation for all of these materials. We also confirm that the computed properties are in good agreement with corresponding theoretical results from literature.

In the second test group, we randomly select 50 binary and 50 ternary compounds from Inorganic Crystal Structure Database



**Table 1**

The list of sub-directories and their descriptions. If LDA potential is used, GGA in directory names changes to LDA.

| Directory         | Description   |
|-------------------|---|
| INPUT0            | Input files for VASP calculation <ul style="list-style-type: none"> <li>– POSCAR_rlx_GGA: optimized structure</li> <li>– KPOINTS: converged <b>k</b>-point grid</li> <li>– INCAR: VASP input file with converged cutoff energy and magnetic spin ordering</li> </ul>  |
| kptest            | <b>k</b> -point convergence test <ul style="list-style-type: none"> <li>– kpoint.log: calculation log for <b>k</b>-points</li> </ul>  |
| encut             | Cutoff energy convergence test <ul style="list-style-type: none"> <li>– cutoff.log: calculation log for cutoff energy</li> </ul>  |
| magnetic_ordering | Identifying magnetic spin ordering  |
| relax_GGA         | Structure relaxation  |
| band_GGA          | Band structure and band gap   |
| dos_GGA           | Density of states   |
| dielectric_GGA    | Dielectric constant   |
| hybrid_GGA        | Band-gap calculation with HSE@PBE scheme  |
| effm_GGA          | Effective mass  |
| Results           | There are two sub-directories, electron and hole. Main computational results <ul style="list-style-type: none"> <li>– POSCAR_GGA: optimized structure</li> <li>– Band_gap_GGA.log: information of band gap</li> <li>– band_GGA.png (band_GGA.pdf): band structure image.</li> <li>– band_corrected.png (band_corrected.pdf): image of gap-corrected band structure.</li> <li>– Band_gap_hybrid_GGA.log: band gap with HSE@PBE scheme.</li> <li>– dos_GGA.png (dos_GGA.pdf): image of density of states.</li> <li>– dielectric_GGA.log: information on dielectric constant.</li> <li>– effective_mass_hole_GGA.log: effective mass of hole.</li> <li>– effective_mass_electron_GGA.log: effective mass of electron.</li> <li>– Properties.json: summary of material properties.</li> </ul> |
| INPUT0_old        | Input files for calculation with ferromagnetic ordering. If a more stable spin ordering is found, this directory is created.  |
| relax_GGA_old     | Structure relaxation with ferromagnetic ordering. If a more stable spin ordering is found, this directory is created.   |
| name_with_U       | AMP <sup>2</sup> calculation with DFT+ <i>U</i> calculation. If the material is metallic and DFT+ <i>U</i> calculation has been conducted, all the results move to this directory.  |

(ICSD) [48] under a condition that no elements appear in more than five compounds. The list of materials and success/failure are compiled in Table S1. In summary, we find that AMP<sup>2</sup> successfully finishes for 86 materials. Among the failed 14 cases are three crashes due to insufficient memories and three failed convergences in electronic steps during **k**-point tests. The most frequent failures (8 cases) occurred in identifying magnetic ordering. Specifically, during supercell calculations to obtain exchange parameters in Ising models, electronic steps could not converge sufficiently. (AMP<sup>2</sup> tried to find a magnetic ordering in 16 compounds in total with 8 successes and 8 failures.) In our opinion, this is mainly because the designated magnetic ordering is metastable. Coupled with inherent slow convergences in spin-polarized calculations, the metastability results in a tricky convergence behavior. We note that when AMP<sup>2</sup> fails in finding a stable magnetic ordering, it still proceeds by assuming a ferromagnetic ordering, leaving a warning message in the log file.

There are limitations of AMP<sup>2</sup> due to specific computational setups. For instance, the package uses pseudopotentials for lanthanides that treat *f* orbitals as core levels. When users employ pseudopotentials that include *f* as valence levels, it is recommended to apply *U* term to *f* orbitals [49]. In the case of Ti ions,

AMP<sup>2</sup> does not apply a PBE + *U* method since *d* levels are empty in Ti<sup>4+</sup>, the most frequent valence state of Ti ions. This means that electronic structures and magnetic ordering of materials with Ti<sup>2+</sup> and Ti<sup>3+</sup> (for example, TiO and LaTiO<sub>3</sub>, respectively) could be incorrect. In addition, the present version does not take into account the van der Waals interaction so it may produce incorrect crystal structures for layered materials such as MoS<sub>2</sub>. The fractional occupation of ionic sites is also not supported yet [50].

In Section 3.1, we used convergence criteria for stress tensors as 10 kbar. In soft materials with small elastic constants such as organic crystals, errors in structural parameters can be substantial. For instance, optimized lattice parameters of a C<sub>3</sub>H<sub>4</sub>N<sub>4</sub> crystal differ from the results with current setting by more than 10% when tighter convergence conditions of 1 meV/atom and 1 kbar are employed. Therefore, materials with small elastic moduli require more strict conditions. (Detailed instructions are given in the manual.) In a future update, we will adjust the convergence threshold for stress tensors according to preliminary estimation of elastic constants.

## 5. Conclusion

To summarize, we introduced AMP<sup>2</sup> targeted for fully automating DFT calculations with minimal user interferences. It is confirmed that AMP<sup>2</sup> provides the band structure, band gap, effective mass, density of states and dielectric constant that are highly reliable. Notably, the package finds the antiferromagnetic ground state and addresses band-gap underestimations in semilocal functionals with the help of a hybrid functional.

There are several databases that provide theoretical properties for most inorganic materials present in ICSD [51–53]. Nevertheless, we believe that the present package is useful because the current databases lack in key features of AMP<sup>2</sup>, in particular addressing band-gap underestimation and identifying an antiferromagnetic ordering, which can result in critical errors in the database. For instance, the databases as of now records that Ge is metal. Another example is NiO that is calculated only in the ferromagnetic ordering with the band gap of 2.3 eV [51]. AMP<sup>2</sup> correctly identifies the antiferromagnetic ground state [54] with the band gap of 4.5 eV, which is in good agreement with experimental value of 4.3 eV [55]. Therefore, users may utilize AMP<sup>2</sup> in obtaining more accurate properties than for the current databases.

In conclusion, we believe that AMP<sup>2</sup> will significantly expand the usage of DFT calculations by making the computation more accessible. In addition, since AMP<sup>2</sup> is available in an open-source platform, the package may serve to consolidate technical know-hows from various research groups into a single package. Even though the present package focuses on the crystalline properties, future releases will include more advanced features such as the calculation on defects and surfaces.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This work was supported by Creative Materials Discovery Program (RIAM) through the National Research Foundation of Korea (NRF) funded by Ministry of Science and ICT (2017M3D1A1040689). The computations were carried out at the KISTI supercomputing center (KSC-2019-CRE-0131).



## Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.cpc.2020.107450>.

## References

- [1] W. Kohn, L.J. Sham, *Phys. Rev.* 140 (1965) A1133–A1138.
- [2] J.K. Nørskov, T. Bligaard, J. Rossmeisl, C.H. Christensen, *Nature Chem.* 1 (2009) 37–46.
- [3] S. Curtarolo, G.L.W. Hart, M.B. Nardelli, N. Mingo, S. Sanvito, O. Levy, *Nature Mater.* 12 (2013) 191–201.
- [4] K. Yim, Y. Yong, J. Lee, K. Lee, H.-H. Nahm, J. Yoo, C. Lee, C. Seong Hwang, S. Han, *NPG Asia Mater.* 7 (2015) e190.
- [5] M. Lee, Y. Youn, K. Yim, S. Han, *Sci. Rep.* 8 (2018) 14794.
- [6] M. Aykol, S. Kim, V.I. Hegde, D. Snyder, Z. Lu, S. Hao, S. Kirklin, D. Morgan, C. Wolverton, *Nature Commun.* 7 (2016) 13779.
- [7] G. Hautier, A. Miglio, G. Ceder, G.-M. Rignanese, X. Gonze, *Nature Commun.* 4 (2013) 2292.
- [8] K. Yim, Y. Youn, M. Lee, D. Yoo, J. Lee, S.H. Cho, S. Han, *Npj Comput. Mater.* 4 (2018) 17.
- [9] Y. Youn, M. Lee, D. Kim, J.K. Jeong, Y. Kang, S. Han, *Chem. Mater.* 31 (2019) 5475–5483.
- [10] Y. Kang, Y. Youn, S. Han, J. Park, C.-S. Oh, *Chem. Mater.* 31 (2019) 4072–4080.
- [11] T. Muller, G. Hautier, A. Jain, G. Ceder, *Chem. Mater.* 23 (2011) 3854–3862.
- [12] S. Körbel, M.A.L. Marques, S. Botti, *J. Mater. Chem. C* 4 (2016) 3157–3167.
- [13] A. Jain, Y. Shin, K.A. Persson, *Nat. Rev. Mater.* 1 (2016) 15004.
- [14] G. Kresse, *Phys. Rev. B* 54 (1996) 11169–11186.
- [15] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M.B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A.D. Corso, S. de Gironcoli, P. Delugas, R.A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N.L. Nguyen, H.-V. Nguyen, A. Otero-de-la Roza, L. Paulatto, S. Poncè, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A.P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, S. Baroni, *J. Phys. Condens. Matter.* 29 (2017) 465901.
- [16] J.M. Soler, E. Artacho, J.D. Gale, A. García, J. Junquera, P. Ordejón, D. Sánchez-Portal, *J. Phys. Condens. Matter.* 14 (2002) 2745–2779.
- [17] X. Gonze, F. Jollet, F. Abreu Araujo, D. Adams, B. Amadon, T. Applencourt, C. Audouze, J.-M. Beuken, J. Bieder, A. Bokhanchuk, E. Bousquet, F. Bruneval, D. Caliste, M. Côté, F. Dahm, F. Da Pieve, M. Delaveau, M. Di Gennaro, B. Dorado, C. Espejo, G. Geneste, L. Genovese, A. Gerossier, M. Giantomassi, Y. Gillet, D.R. Hamann, L. He, G. Jomard, J. Laflamme Janssen, S. Le Roux, A. Levitt, A. Lherbier, F. Liu, I. Lukačević, A. Martin, C. Martins, M.J.T. Oliveira, S. Poncè, Y. Pouillon, T. Rangel, G.-M. Rignanese, A.H. Romero, B. Rousseau, O. Rubel, A.A. Shukri, M. Stankovski, M. Torrent, M.J. Van Setten, B. Van Troeye, M.J. Verstraete, D. Waroquiers, J. Wiktor, B. Xu, A. Zhou, J.W. Zwanziger, *Comput. Phys. Comm.* 205 (2016) 106–131.
- [18] S.J. Clark, M.D. Segall, C.J. Pickard, P.J. Hasnip, M.I.J. Probert, K. Refson, M.C. Payne, *Z. Krist.* 220 (2009) 567–570.
- [19] G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari, B. Kozinsky, *Comput. Mater. Sci.* 111 (2016) 218–230.
- [20] K. Mathew, J.H. Montoya, A. Faghaninia, S. Dwarakanath, M. Aykol, H. Tang, I. Chu, T. Smidt, B. Bocklund, M. Horton, J. Dagdelen, B. Wood, Z.-K. Liu, J. Neaton, S.P. Ong, K. Persson, A. Jain, *Comput. Mater. Sci.* 139 (2017) 140–152.
- [21] S.P. Ong, W.D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V.L. Chevrier, K.A. Persson, G. Ceder, *Comput. Mater. Sci.* 68 (2013) 314–319.
- [22] W. Setyawan, S. Curtarolo, *Comput. Mater. Sci.* 49 (2010) 299–312.
- [23] J.P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* 77 (1996) 3865–3868.
- [24] J.P. Perdew, A. Zunger, *Phys. Rev. B* 23 (1981) 5048–5079.
- [25] J. Heyd, G.E. Scuseria, M. Ernzerhof, *J. Chem. Phys.* 118 (2003) 8207–8215.
- [26] AMP<sup>2</sup> manual can be found at the Supplementary data.
- [27] S.L. Dudarev, G.A. Botton, S.Y. Savrasov, C.J. Humphreys, A.P. Sutton, *Phys. Rev. B* 57 (1998) 1505–1509.
- [28] L. Wang, T. Maxisch, G. Ceder, *Phys. Rev. B* 73 (2006) 195107.
- [29] X. Yu, C.-F. Huo, Y.-W. Li, J. Wang, H. Jiao, *Surf. Sci.* 606 (2012) 872–879.
- [30] X.-Q. Chen, R. Podloucky, P. Rogl, *J. Appl. Phys.* 100 (2006) 113901.
- [31] J. Ihm, A. Zunger, M.L. Cohen, *J. Phys. C Solid State Phys.* 12 (1979) 4409–4422.
- [32] H.J. Monkhorst, J.D. Pack, *Phys. Rev. B* 13 (1976) 5188–5192.
- [33] VASP manual. [https://cms.mpi.univie.ac.at/wiki/index.php/AMIX\\_MAG](https://cms.mpi.univie.ac.at/wiki/index.php/AMIX_MAG) (accessed 22 August 2019).
- [34] K. Lee, Y. Youn, S. Han, *Sci. Technol. Adv. Mater.* 18 (2017) 246–252.
- [35] A. Togo, I. Tanaka, *Scr. Mater.* 108 (2015) 1–5.
- [36] P.E. Blöchl, O. Jepsen, O.K. Andersen, *Phys. Rev. B* 49 (1994) 16223–16233.
- [37] S. Park, B. Lee, S.H. Jeon, S. Han, *Curr. Appl. Phys.* 11 (2011) S337–S340.
- [38] J.P. Perdew, M. Ernzerhof, K. Burke, *J. Chem. Phys.* 105 (1996) 9982–9985.
- [39] M.A.L. Marques, J. Vidal, M.J.T. Oliveira, L. Reining, S. Botti, *Phys. Rev. B* 83 (2011) 035119.
- [40] C. Kittel, *Introduction to Solid State Physics*, sixth ed., Wiley, New York, 1986.
- [41] M.P. Dare-Edwards, J.B. Goodenough, A. Hamnett, A. Katty, *Mater. Res. Bull.* 19 (1984) 435–442.
- [42] L.H. Tjeng, M.B.J. Meinders, J. van Elp, J. Ghijsen, G.A. Sawatzky, R.L. Johnson, *Phys. Rev. B* 41 (1990) 3190–3199.
- [43] K.H. Kim, C.U. Jung, T.W. Noh, S.C. Kim, *Phys. Rev. B* 55 (1997) 15393–15396.
- [44] G.K.H. Madsen, J. Carrete, M.J. Verstraete, *Comput. Phys. Comm.* 231 (2018) 140–145.
- [45] M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller, F. Bechstedt, *Phys. Rev. B* 73 (2006) 045112.
- [46] T. Hahn, U. Shmueli, A.J.C. Wilson, *International Union of Crystallography (Eds.), International Tables for Crystallography*, D. Reidel Pub. Co., 1984, Sold and distributed in the U.S.A. and Canada by Kluwer Academic Publishers Group, Dordrecht, Holland ; Boston, U.S.A. : Hingham, MA.
- [47] SNUMAT website. <https://www.snumat.com>.
- [48] G. Bergerhoff, R. Hundt, R. Sievers, I.D. Brown, *J. Chem. Inf. Comput. Sci.* 23 (1983) 66–69.
- [49] P. Larson, W.R.L. Lambrecht, A. Chantis, M. van Schilfgaarde, *Phys. Rev. B* 75 (2007) 045114.
- [50] K. Yang, C. Oses, S. Curtarolo, *Chem. Mater.* 28 (2016) 6484–6492.
- [51] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K.A. Persson, *APL Mater.* 1 (2013) 011002.
- [52] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R.H. Taylor, L.J. Nelson, G.L.W. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, O. Levy, *Comput. Mater. Sci.* 58 (2012) 227–235.
- [53] S. Kirklin, J.E. Saal, B. Meredig, A. Thompson, J.W. Doak, M. Aykol, S. Rühl, C. Wolverton, *Npj Comput. Mater.* 1 (2015) 15010.
- [54] S. Park, H.-S. Ahn, C.-K. Lee, H. Kim, H. Jin, H.-S. Lee, S. Seo, J. Yu, S. Han, *Phys. Rev. B* 77 (2008) 134103.
- [55] G.A. Sawatzky, J.W. Allen, *Phys. Rev. Lett.* 53 (1984) 2339–2342.