

ARTICLE OPEN



Exchange-correlation functionals for band gaps of solids: benchmark, reparametrization and machine learning

Pedro Borlido¹, Jonathan Schmidt², Ahmad W. Huran², Fabien Tran³, Miguel A. L. Marques²✉ and Silvana Botti¹✉

We conducted a large-scale density-functional theory study on the influence of the exchange-correlation functional in the calculation of electronic band gaps of solids. First, we use the large materials data set that we have recently proposed to benchmark 21 different functionals, with a particular focus on approximations of the meta-generalized-gradient family. Combining these data with the results for 12 functionals in our previous work, we can analyze in detail the characteristics of each approximation and identify its strong and/or weak points. Beside confirming that mBJ, HLE16 and HSE06 are the most accurate functionals for band gap calculations, we reveal several other interesting functionals, chief among which are the local Slater potential approximation, the GGA AK13LDA, and the meta-GGAs HLE17 and TASK. We also compare the computational efficiency of these different approximations. Relying on these data, we investigate the potential for improvement of a promising subset of functionals by varying their internal parameters. The identified optimal parameters yield a family of functionals fitted for the calculation of band gaps. Finally, we demonstrate how to train machine learning models for accurate band gap prediction, using as input structural and composition data, as well as approximate band gaps obtained from density-functional theory.

npj Computational Materials (2020)6:96; <https://doi.org/10.1038/s41524-020-00360-0>

INTRODUCTION

Over the past decades, density-functional theory (DFT)^{1,2} has become the workhorse theory in computational chemistry and solid-state physics. This powerful approach is an exact and elegant reformulation of the many-body quantum mechanics that governs the behavior of electrons in all kinds of systems (atom, molecule, solid, etc.). Moreover, the Kohn–Sham equations that stem from the theory² can be solved efficiently with modern computers. These equations rely on a single approximation, namely the one for the exchange–correlation (xc) energy, which is responsible for the accuracy of the calculations³. A very large number, perhaps more than 500^{4,5}, of such approximations have appeared in the literature over the last 50 years. However, only a handful of them, such as the PBE from Perdew, Burke, and Ernzerhof^{6,7} or the hybrids B3LYP^{8,9} and HSE06 from Heyd, Scuseria, and Ernzerhof^{10,11}, have found widespread use.

In principle, the exact xc functional is universal, i.e., it can be used for any system of electrons interacting through the Coulomb potential and for any (ground-state) property. Some approximate functionals however, are designed to provide a particular good description of a specific quantity, such as for instance the nuclear magnetic resonance chemical shift¹², ionization potential¹³, formation energy¹⁴, or electronic band gap^{15–17}. Such approach favors the best description of the target property, at the cost of an accompanying reduction of the scope of the functional. Of course, for many specific applications, accuracy is more important than universality, so this path is widely approved.

A large amount of attention has been devoted by the computational solid-state community to the problem of predicting the fundamental band gap of materials. This is due to the relevance of the band gap in technological applications such as opto-electronics and photovoltaics^{18,19}.

A proper evaluation of the band gap is one of the greatest challenges for Kohn–Sham DFT. The fundamental gap (E_G) for a system of N electrons is defined as the difference between its ionization potential (I) and its electron affinity (A),

$$E_G = I - A, \quad (1)$$

meaning that it can be expressed as a difference of ground-state total energies, which are accessible to DFT. On the other hand, the Kohn–Sham (KS) band gap is defined as the difference between the eigenvalues of the conduction band minimum (CBM) and the valence band maximum (VBM):

$$E_g^{KS} = \epsilon_{CBM} - \epsilon_{VBM}. \quad (2)$$

In exact Kohn–Sham DFT these two quantities are not equal, as they differ by the so-called derivative discontinuity Δ_{xc} ^{20,21}:

$$E_G = E_g^{KS} + \Delta_{xc}. \quad (3)$$

The exchange–correlation functionals that depend explicitly on the electron density, such as local density approximations (LDAs) and generalized gradient approximations (GGAs), have a derivative discontinuity that is exactly zero ($\Delta_{xc}^{LDA, GGA} = 0$) for solids²². This is often presented as a justification for the systematic underestimation of band gaps in these approximations^{21,23,24}. More sophisticated functionals, like hybrids and meta-GGAs, which depend on the orbitals, can possess a non-vanishing derivative discontinuity Δ_{xc} . Therefore, within the standard Kohn–Sham theory, a theoretical estimation of the band gap must include both terms, as in Eq. (3).

However, hybrid and meta-GGA functionals are often not implemented in the standard Kohn–Sham formalism, where the Kohn–Sham potential is strictly local, since this would require using the optimized effective potential method^{25–27}, i.e. solving a complicated integral equation. Instead, the large majority of

¹Institut für Festkörpertheorie und -optik, Friedrich-Schiller-Universität Jena and European Theoretical Spectroscopy Facility, Max-Wien-Platz 1, 07743 Jena, Germany. ²Institut für Physik, Martin-Luther-Universität Halle-Wittenberg, 06099 Halle, Germany. ³Institute of Materials Chemistry, Vienna University of Technology, Getreidemarkt 9/165-TC, 1060 Vienna, Austria. ✉email: miguel.marques@physik.uni-halle.de; silvana.botti@uni-jena.de

codes^{28–32} implements meta-GGAs and hybrids within the generalized Kohn–Sham (GKS) formalism³³, which leads to a non-local potential. In this formalism, it was shown that the generalized Kohn–Sham gap E_g^{GKS} is a direct approximation to E_G and that no derivative discontinuity should be added^{34–38}. This is also consistent with the interpretation of hybrid functionals as approximations to the self-energy of the many-body *GW* theory³⁹. Comparing the generalized Kohn–Sham band gap of a hybrid or meta-GGA functional to the experimental photoemission gap is hence justified. Note that for explicit functionals of the electron density, the Kohn–Sham and generalized Kohn–Sham formalisms are equivalent and lead to the same (semilocal) potential. For a deeper discussion on the definition of band gaps of solids, we refer the reader to ref. ³⁴.

Of course, most of the experimentally measured band gaps are optical band gaps, which differ from the photoemission band gap by the excitonic binding energy. The latter can, in principle, be calculated by solving the many-body Bethe–Salpeter equation⁴⁰ (or by using time-dependent DFT⁴¹). However, in practice, for the majority of bulk materials the excitonic binding energy amounts to tens of meV⁴², a much smaller quantity than typical errors of DFT functionals.

Recently, we developed a large material data set with the objective of benchmarking approximate xc functionals for the calculation of band gaps⁴³. This data set includes a large variety of semiconductors and insulators, with small, intermediate, and large band gaps. Furthermore, it includes materials with chemical elements that span almost the complete periodic table. We performed DFT calculations for all these materials using 12 functionals of the LDA, GGA, meta-GGA, and hybrid types⁴³. The most accurate functional turned out to be the popular modified (m) Becke–Johnson (BJ) with LDA correlation (mBJLDA) meta-GGA potential¹⁵, followed very closely by the GGA high-local exchange (HLE16)¹⁶ and the screened hybrid HSE06^{10,11,44}.

Here, we will take a step further: we investigate the performance of 21 additional xc functionals to obtain a complete picture of the advantages and limitations of each approximation. Building on this information, we propose then reparametrizations targeted to improve the accuracy of the most reliable functionals in the determination of band gaps.

Most of the functionals that we study in this work are modern meta-GGAs, which should potentially be more accurate than LDAs and standard GGAs for the calculation of band gaps. Actually, a few of them were developed specifically for band gaps (such as the meta-GGA HLE17¹⁷), whereas others were developed for better accuracy in general. Besides meta-GGAs, less studied LDA- and GGA-type functionals are also considered.

Our choice to benchmark DFT functionals leaves out several other approaches proposed in the literature to tackle the prediction of band gaps. These cover methods like DFT+U^{45,46}, DFT–1/2^{47–49}, Koopmans-compliant functionals⁵⁰, dielectric-dependent hybrids^{51,52}, and self-interaction correction methods⁵³. Their inclusion, which we leave for future work, would require additional computational resources that are not available.

All xc functionals depend in some way on internal parameters that are either determined from theoretical considerations or by fitting accurate ab initio results or experimental data. In order to design a new and improved approximation it is essential to understand the behavior of the chosen analytical form with respect to variations of its internal parameters. As such, we focus in this work on the error in the band gaps as a function of the internal parameters of selected functional forms. This analysis allows us also to select improved empirical parametrizations by minimizing the average error within a data set.

Finally, we show how to use this large data set to train a machine learning model for band gap prediction. In fact, band gaps calculated using different functionals can be used as “estimators”. In principle, it should be possible to decrease the

error in the calculation of band gaps by combining several estimators using statistical methods. We decided therefore to experiment this strategy testing few machine learning models that are especially adequate for this task⁵⁴. There has already been a number of machine learning studies concerning the prediction of band gaps. One research direction is the direct prediction of DFT band gaps to avoid ab initio calculations^{55,56}. However, this approach provides at best an accuracy that is slightly worse than the accuracy of the functional used in generating the training data. A second approach relies on learning higher fidelity theoretical band gaps, like in ref. ⁵⁷, which leads to a slightly improved accuracy in comparison to experimental data. Finally, one can directly learn experimental band gaps, as in ref. ⁵⁸. We will follow the latter approach, combining it with the use of various functionals as input features for the machine learning model, to achieve higher accuracy than traditional DFT calculations.

The remainder of this paper is organized as follows. In the following, we provide a brief overview of the tested functionals, followed by a description of the large-scale benchmark and its results in “Results and discussion”. In the “Functional reparametrization” section, we study the behavior of selected functionals as a function of their internal parameters, which leads to the new parametrizations that are then tested and analyzed. In the “Machine learning” section, we investigate machine learning models that use our data to improve the estimation of band gaps. We finish with a summary and conclusions. Details on our calculations are given in the “Methods” section.

The 12 xc functionals/potentials that were already tested in our previous work⁴³ are the LDA PZ81⁵⁹, three GGAs (PBE^{6,7}, PBEsol⁶⁰, and HLE16¹⁶), three meta-GGAs (BJ⁶¹, mBJLDA¹⁵, that we call here for simplicity mBJ, and SCAN⁶²), three traditional hybrids (HSE06^{10,11,44}, HSE14⁶³, and PBE0^{64,65}), and two hybrids with a system-dependent mixing parameter (HSE_{mix}³⁹ and PBE0_{mix}³⁹). A description of these functionals can be found in ref. ⁴³. In the following, we provide a brief description of the additional functionals selected for the present work. The materials database is also presented in ref. ⁴³.

We start at the LDA level of approximation, i.e., first rung of Jacob’s ladder⁶⁶, considering the local Slater potential (Sloc)⁶⁷. Its starting point is some general, yet very simple, form of LDA-type exchange potential:

$$v_x = -an^b, \quad (4)$$

where $n = \sum_j^{\text{occ}} |\psi_j|^2$, built with the Kohn–Sham orbitals ψ_j , is the electron density. This functional form, that encompasses both the LDA⁶⁸ and the Slater *Xa* method⁶⁹, was fitted to the non-local Slater potential⁷⁰, calculated with the Hartree–Fock orbitals⁷¹ of a series of closed-shell atoms (Be, Ne, Mg, Ar, Ca, Zn, Kr, Sr, Pd, and Xe), to determine the fitting parameters $a = 1.67$ and $b = 0.3$. It is obvious that Sloc violates several exact conditions, most strikingly the requirement of uniform coordinate scaling of exchange, which requires $b = 1/3$ ⁷². Furthermore, and despite the fact that this functional is technically a LDA, it does not recover the exact exchange energy of the homogeneous electron gas for uniform densities, which is obtained only if $a = (3/\pi)^{1/3} \approx 0.985$ and $b = 1/3$. Note that the exchange energy functional corresponding to Eq. (4) is $E_x = -a/(b+1) \int n^{b+1} d^3r$.

Moving to the GGA family of functionals, the second rung of Jacob’s ladder, we find the Engel–Vosko exchange functional (EV93)⁷³. Its enhancement factor F_x , which multiplies the LDA-like part of the functional, has a Padé form:

$$F_x^{\text{EV93}} = \frac{1 + a_1 s^2 + a_2 s^4 + a_3 s^6}{1 + b_1 s^2 + b_2 s^4 + b_3 s^6}, \quad (5)$$

where $s = |\nabla n|/(2nk_F)$ is the reduced gradient and $k_F = (3\pi^2 n)^{1/3}$ is the Fermi momentum. In order to recover the gradient expansion, the difference $a_1 - b_1 = 10/81$ is fixed, reducing to

five the number of free-parameters. These were fitted to the local part of the virial relation⁷², obtained from results of the optimized potential method^{25,74} for a set of 20 spherical atoms (Kr, Rb, Sr, Mo, Tc, Pd, Ag, Cd, Sb, Xe, Cs, Ba, Eu, Yb, Re, Pt, Au, Hg, Bi, and Rn). These atoms were chosen because they present extensive regions for which $s < 1$. As EV93 is an exchange-only functional, in this work, as in previous works^{75,76}, it is used in conjunction with the GGA correlation PW91 (EV93PW91). Tested in previous works^{75–77}, EV93PW91 moderately improves over the PBE for band gaps.

The Armiento–Mattson (AM05) xc functional⁷⁸ was constructed in order to reproduce both the homogeneous electron gas and the Airy gas model of the edge electron gas⁷⁹. Similarly to the GGA PBEsol, AM05 was shown to be more accurate than PBE for the lattice constant of solids⁸⁰.

The second-order GGA proposed in 2011 (SOGGA11)⁸¹ was designed, as the name indicates, with the intent of satisfying the second-order gradient expansion, while preserving good energetics for molecules. This GGA defines an exchange enhancement factor F_x that is the sum of two geometric series, whose arguments are in turn the exchange enhancement factors of PBE^{6,7} and RPBE⁸². SOGGA11 treats correlation in a similar fashion, although using inspiration from the GGA of Perdew–Wang (PW91)^{83–85}. Due to its construction, this functional has many parameters (24 in total), most of them fitted to a series of databases. According to the results in ref. ⁸⁶, SOGGA11 is slightly better than the PBE for the band gap of simple semiconductors.

The Armiento–Kümmel functional (AK13)⁸⁷ is an exchange-only approximation designed to mimic the asymptotic behavior ($\propto 1/r$) of finite and semi-infinite (surface) systems, but also to have a reasonable behavior for small s . The authors thus obtained two expressions for F_x , adequate in the corresponding regimes, and by combining them the resulting enhancement factor takes the form

$$F_x^{\text{AK13}} = 1 + B_1 s \ln(1 + s) + B_2 s \ln(1 + \ln(1 + s)). \quad (6)$$

The two coefficients are determined by satisfying the gradient expansion at small s , and demanding that at large distance r from the nuclei the potential has the form $1/r$. AK13 was shown to improve over the PBE for band gaps⁸⁸. Note that there are a series of issues related to the asymptotic form of the AK13 functional^{89–91}; the situation is similar to the one of LB94 and (m)BJ, that are described below. These problems limit considerably the possibility to use the functional for finite systems. However, they are not relevant for the present study, since we are dealing exclusively with infinite periodic systems. As AK13 is an exchange-only functional, we use it together with the LDA correlation PW92⁹². In the following, we denote the combination of AK13 exchange and LDA correlation as AK13LDA.

With the vast majority of GGA approximations to the xc energy the corresponding potential v_{xc} decays too fast, i.e., typically exponentially instead of decaying as $\propto 1/r$ (AK13 is one exception). Van Leeuwen and Baerends⁹³ proposed a solution to this problem by modeling directly the exchange potential so that it obeys the correct asymptotic behavior $\propto 1/r$. Their potential LB94, which is inspired by a previous work of Becke⁹⁴, has a rather simple form:

$$v_{xc}^{\text{LB94}} = v_{xc}^{\text{LDA}} - \left(\frac{n}{2}\right)^{1/3} \frac{\beta x^2}{1 + 3\beta x \operatorname{arcsinh}(x)}, \quad (7)$$

where $x = 2^{1/3} |\nabla n| / n^{4/3} = 2(6\pi^2)^{1/3} s$. As in the case of basically any direct approximation to the xc potential^{74,95}, Eq. (7) is not a functional derivative of an energy functional, and thus violates the zero-force and zero-torque theorems, along with several other exact conditions^{95,96}. On the other hand, its correct asymptotic behavior makes LB94 useful in situations where ionization is important^{97–99}. As indicated in Eq. (7) the LDA correlation PW92 is used.

We turn now to meta-GGA functionals, which depend also on the kinetic-energy density $\tau = (1/2) \sum_j^{\text{occ}} \nabla \psi_j^* \cdot \nabla \psi_j$ and represent the third rung of Jacob’s ladder. The Perdew–Kurth–Zupan–Blaha (PKZB)^{100,101} functional, one of the earliest meta-GGAs, has an

exchange enhancement factor similar to the PBE and recovers the fourth-order gradient expansion. In addition, its correlation component is exactly zero for one-electron densities, i.e., the functional is self-interaction-error free. PKZB, which has two fitted parameters, yields good surface and atomization energies¹⁰⁰, but it provides a rather poor description of the geometry of molecules¹⁰² and solids¹⁰³, as well as hydrogen-bonded systems¹⁰⁴.

The non-empirical meta-GGA Tao–Perdew–Staroverov–Scuseria (TPSS)^{105,106} is built on top of the PKZB functional, and it is forced to satisfy additional exact mathematical constraints. In terms of performance (geometry and energetics), TPSS represents an improvement with respect to PKZB^{105,106}.

Several other functionals were built starting from the TPSS, such as the revised TPSS (revTPSS)^{107,108}. This functional attempts to correct the deficiencies of TPSS for solids (e.g., too large lattice constants), while maintaining the same accuracy for molecular systems. The differences between TPSS and revTPSS concern the reparametrization of several coefficients, to better describe the small gradient expansion, and the change of an exponent in the enhancement factor. This functional yields results close to those of PBEsol⁶⁰, but it suffers from some analytic complications related to its asymptotic behavior¹⁰⁹.

The meta-GGA “made very simple” (MVS)¹¹⁰ is an exchange-only functional, which has a relatively simple form

$$F_x^{\text{MVS}} = \frac{1 + k_0 f_x(a)}{(1 + b s^4)^{1/8}}, \quad (8)$$

where

$$f_x(a) = \frac{1 - a}{\left[(1 + e_1 a^2)^2 + c_1 a^4\right]^{1/4}} \quad (9)$$

and $a = (\tau - \tau^W) / \tau^{\text{TF}}$ is an iso-orbital indicator. In the latter expression, $\tau^W = |\nabla n|^2 / (8n)$ and $\tau^{\text{TF}} = (3/10)(3\pi^2)^{2/3} n^{5/3}$ are the von Weizsäcker¹¹¹ and Thomas-Fermi^{112,113} kinetic-energy densities, respectively. Equation (9) interpolates between the homogeneous electron gas ($a = 1$) and the lower bound¹¹⁴ ($a = 0$) regimes. The constants k_0 , b , e_1 , and c_1 were chosen such that this functional obeys exact constraints, among which the second-order gradient expansion and the large- Z asymptotic expansion of the exchange energy for neutral atoms¹¹⁵. We use MVS exchange in conjunction with the regularized TPSS¹⁰⁷ correlation, sometimes also called vPBE in the literature.

The recent “strongly constrained and appropriately normed” (SCAN)⁶² functional is currently one of the most popular meta-GGAs for the geometry optimization and energetics of solids^{103,116}. It was constructed non-empirically to satisfy all 17 mathematical conditions that are possible to impose to a meta-GGA functional. Several modifications of SCAN have been proposed^{117,118}. One of these is the regularized SCAN (rSCAN)¹¹⁸, which tries to solve some of the numerical problems of SCAN, related to the iso-orbital indicator a defined above. The derivative of a , a common ingredient in the formulation of meta-GGAs, can lead to divergences in some physical situations¹¹⁹, which in turn cause divergences in the xc potential. rSCAN replaces the kinetic energy of the homogeneous electron gas, the iso-orbital indicator and the switching function of the original SCAN, with modified versions that greatly minimize the divergence problems.

Tao and Mo (TM)¹²⁰ proposed an exchange enhancement factor based on a weighted average of the density-matrix expansion (DME) and a fourth-order gradient correction (slowly varying correction, SC):

$$F_x^{\text{TM}} = w F_x^{\text{DME}} + (1 - w) F_x^{\text{SC}}. \quad (10)$$

The explicit form of these terms can be found in ref. ¹²⁰. As for the correlation, the TM functional is based on TPSS correlation, but

with modifications in order, for instance, to improve the description of the low-density region.

Recently, a revised version of TM (revTM)¹²¹ has been proposed. This functional applies two modifications to TM inspired by the TPSS and revTPSS functionals. First, it generalizes a variable \tilde{q} (involving s and a) in the exchange by including an extra parameter b . The generalized \tilde{q} is the same as the one used in TPSS and revTPSS. Second, it modifies the correlation part by turning the constant β into a function of the Wigner-Seitz radius $r_s = [3/(4\pi n)]^{1/3}$, similarly to what revTPSS does. This allows to recover the correct asymptotic behavior in the low- and high-density regimes.

We also consider a very recent addition to the literature, the TASK functional¹²². This functional was built with the goal of introducing ultra-nonlocality, typical of hybrid functionals, at the meta-GGA level. This is achieved by imposing constraints that enforce a sizable contribution to the derivative discontinuity. Although the resulting functional is not outstanding for what concerns atomization energies of molecules, it was found to give a considerable improvement for band gaps of typical semiconductors with respect to PBE and SCAN¹²².

Turning now to more empirical meta-GGAs, the “high local exchange” proposed in 2017 (HLE17)¹⁷ is a re-scaled version of the TPSS using the same ideas of the GGA HLE16¹⁶. Specifically, the exchange and correlation parts are scaled by 1.25 and 0.5, respectively:

$$E_{xc}^{\text{HLE17}} = 1.25E_x^{\text{TPSS}} + 0.5E_c^{\text{TPSS}}. \quad (11)$$

This empirical scaling was chosen such that HLE17, as HLE16¹⁶, can yield good results for the band gaps of solids and excitation energies of molecules. As reported in ref. ¹⁷, HLE17 performs as well as the hybrid HSE06 for simple semiconductors, but is much worse for antiferromagnetic transition-metal oxides.

The functionals of the Minnesota family are also empirical. Although these functionals belong to several different rungs of Jacob’s ladder, they are sometimes grouped together as they originate from the same research group and follow similar philosophies. These are highly parameterized functionals (often depending on more than 15 parameters) fitted to a series of databases and they aim at a good description of general properties. Among them we will focus on the following meta-GGAs: M06-L^{123,124}, revM06-L¹²⁵, M11-L¹²⁶, MN12-L¹²⁷, and MN15-L¹²⁷.

In contrast to all meta-GGAs presented above, the one of Räsänen, Pittalis, and Proetto (RPP)¹²⁸ is only a potential, with no corresponding energy functional. Actually, it belongs to the well-known family of BJ exchange potential approximations⁶¹. Similarly to BJ, RPP is built as a correction to the Slater potential⁷⁰, that is usually accurately approximated¹²⁹ by the Becke–Roussel (BR) formula¹³⁰, and takes the form

$$v_x^{\text{RPP}} = v_x^{\text{Slater/BR}} + \frac{1}{\pi} \sqrt{\frac{5}{12}} \sqrt{\frac{D}{n}}, \quad (12)$$

where

$$D = 2\tau - \frac{1}{4} \frac{|\nabla n|^2}{n}. \quad (13)$$

The RPP potential, which differs from the original BJ potential by just the extra term $-|\nabla n|^2/(4n)$ in Eq. (13), approaches zero asymptotically, is exact for any single-orbital system (in contrast with BJ) as well as for the homogeneous electron gas. It has been shown to be as accurate as the BJ potential for the ionization potential, electron affinity, and polarizability of finite systems¹³¹. We mention that, as originally proposed¹²⁸, D also contains an extra term depending on the paramagnetic density-current \mathbf{j} , however this term is not relevant here since it is zero in the present context (non-magnetic solids). As in the case of BJ, the

RPP potential is used with the BR term in Eq. (12) and is combined with the LDA correlation potential of PW92⁹².

RESULTS AND DISCUSSION

As a reliable evaluation of performance cannot be reduced to a single value, we will use an assortment of statistically relevant quantities in our analysis¹³². Specifically, we resort to Kendall’s¹³³ rank correlation τ and Pearson’s correlation coefficients r ; the coefficients of the linear fit $y = ax + b$ for the calculated versus experimental band gaps; the mean absolute error $\text{MAE} = \sum_{i=1}^n |y_i - y_{i,\text{exp}}|/n$; the mean error $\text{ME} = \sum_{i=1}^n (y_i - y_{i,\text{exp}})/n$; the error standard deviation $\sigma = \sqrt{\sum_{i=1}^n (y_i - y_{i,\text{exp}} - \text{ME})^2/n}$; the median error MnE ; the interquartile range IQR ; the median of the absolute deviations from the median MADM ; the mean absolute percentage error $\text{MAPE} = 100 \times \sum_{i=1}^n |y_i - y_{i,\text{exp}}|/(n y_{i,\text{exp}})$; and the mean percentage error $\text{MPE} = 100 \times \sum_{i=1}^n (y_i - y_{i,\text{exp}})/(n y_{i,\text{exp}})$.

We emphasize that the experimental values of the band gaps have an error bar that is often very hard to determine. Moreover, there are a myriad of effects (temperature, sample quality, fitting, etc.) that can limit the accuracy of the experimental values. In the context of our data set this was already extensively discussed in ref. ⁴³.

In order to gather more insight into the behavior of the functionals, we further separate our analysis on the subsets of sp and fd materials, where the former consists of materials composed of chemical elements with only s and p valence electrons, and the latter includes the remaining compounds. We consider also the subset, denoted TB67, obtained from intercepting the present data set and that of ref. ⁷⁶. This allows us to compare the effects of data set size on the computed quantities. Apart from the size, the main difference between these two data sets is the absence of magnetic compounds in the present work (six antiferromagnetic oxides were in the data set of 76 compounds used in ref. ⁷⁶).

Ideally, we would generate further subsets referring to the nature of the compounds (ionic, covalent, van der Waals, etc.). Although physical intuition might be sufficient to classify some compounds in this manner, the lack of a rigorous separation criteria means that we cannot do this on the large scale of the present data set.

In order to present the statistical data, we resort to radar plots that provide a rather intuitive overview over the most important statistical quantities that we use to quantify the errors. The exact numerical values are presented in Table S1 of the Supplementary Information.

Error averages for the sets of materials containing each element of the periodic table, along with error averages for specific gaps ranges can be found in Supplementary Figs. S1 and S2. In addition, Supplementary Figs. S3–S26 show a comparison of band-gap values and errors between selected pairs of functionals. Supplementary Fig. 27 compares other statistical variables for the whole dataset and Fig. S28 displays the amount of false metals predicted by each functional. Table SII lists ICSD-ids and compositions of the 85 materials used for the reparametrization of the functionals, while Table SIII reports the compositional features used for machine learning feature selection. The Supplementary Information also contains a spreadsheet with all calculated band gaps.

In Figs. 1–3, we depict the radar plots for all functionals studied here. For comparison, we include also the functionals studied in ref. ⁴³, although in the following we will focus on the new results. The discussion of the results in “LDA”, “GGA”, and “meta-GGA” sections will mainly be based on the analysis of ME, MAE, MPE, and MAPE. A brief discussion on the analysis using the other statistical variables is provided in the “Other metrics” section.

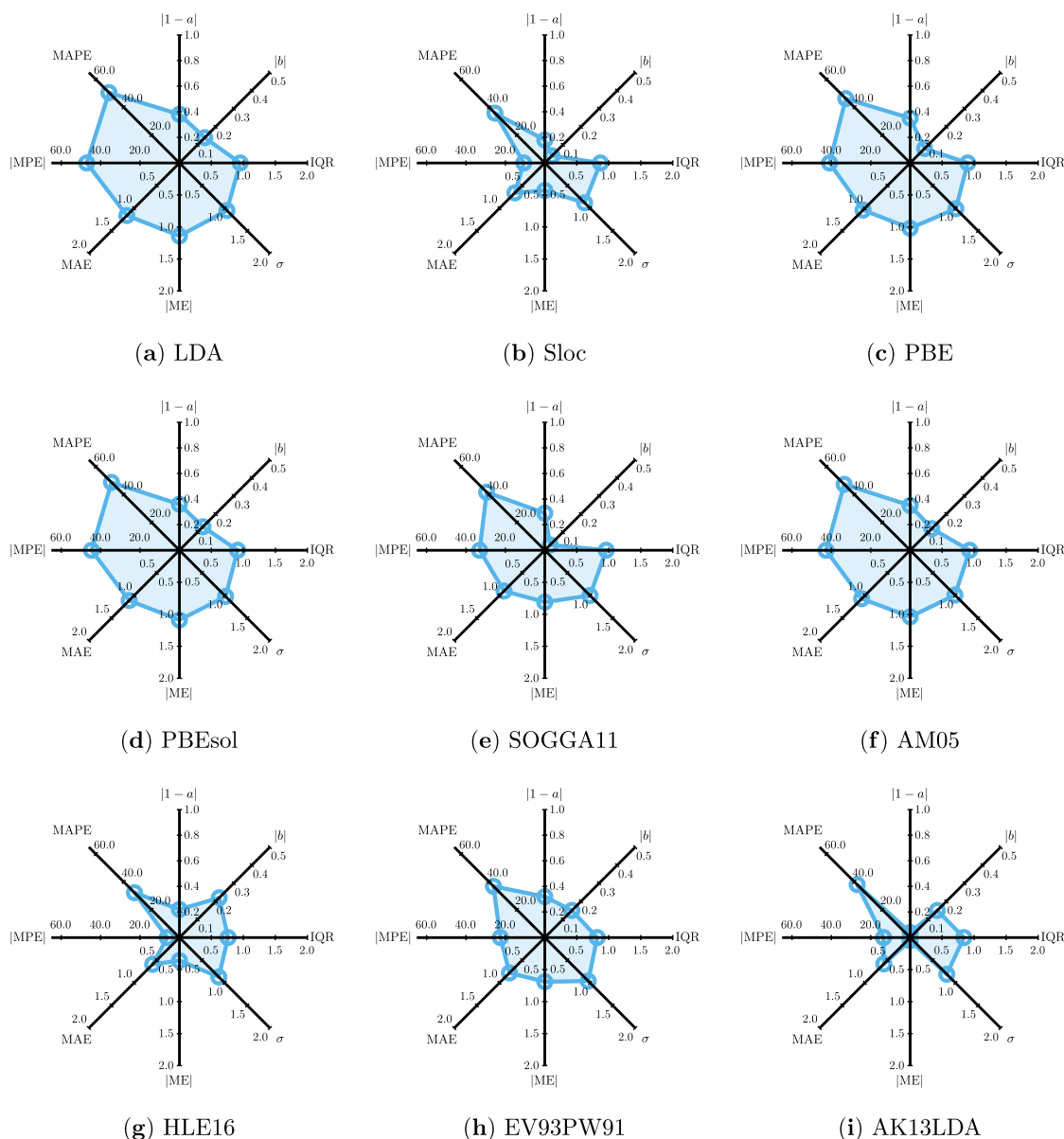


Fig. 1 Radar plots displaying the statistical quantities calculated in this benchmark (MAE, ME, $|b|$, σ , and IQR in eV; MPE and MAPE in %). Here we show the radar plots of the LDA and GGA functionals except LB94.

LDA

We start our analysis of the errors with Sloc (Fig. 1b). Although it is a very simple functional, it presents remarkably small errors. With a MAPE of 36% and a MPE of -11% , it surpasses both LDA (51% and -47% , respectively) and PBE (46% and -41% , respectively) in this front (see also Supplementary Fig. S3). In terms of ME (-0.43 eV) and MAE (0.66 eV), the Sloc errors are about twice smaller than with LDA and PBE. With the smaller sample size of the TB67 subset, the ME and MAE change noticeably (they increase to -0.71 and 0.91 eV, respectively), but to a lesser extent for the MPE and MAPE.

Throughout the periodic table Sloc presents low MAPE, with the exception of Pd and Pb (Supplementary Fig. S1). As usual, the inclusion of spin-orbit coupling should improve the results for solids with the Pb element when Sloc gives a too large band gap.

GGA

At the GGA level, the LB94 potential stands as basically the worst performing method (Fig. 2a), being on almost equal footing with

LDA, PBE, and PBEsol (Supplementary Fig. S4). For the majority of compounds, LB94 increases the band gap with respect to LDA/PBE/PBEsol, but this is moderate and by far not systematic, since in many cases the LB94 band gap is actually smaller. Furthermore, the number of predicted false metals is 50, while it was in the range 30–35 for LDA/PBE/PBEsol⁴³.

The band gap values calculated with AM05 (Fig. 1f) are very similar to the ones obtained with PBE (Supplementary Fig. S5). With 37 false metals, AM05 performs similarly to LDA/PBE/PBEsol and represents a slight improvement with respect to LB94.

The SOGGA11 functional (Fig. 1e) has a better performance than LDA and all GGAs discussed above, as it yields in average smaller errors, as seen from the M(A)E of -0.81 eV (0.90 eV) and M(A)PE of -33% (42%). It predicts 28 false metals, a slightly better result than PBE. In spite of this improvement, the underestimation of the band gaps is still large.

EV93PW91 (Fig. 1h) further improves the results of SOGGA11, with a M(A)E of -0.7 eV (0.8 eV) and M(A)PE -23% (37%). This improvement is seen in both *sp* and *fd* subsets, and in general in

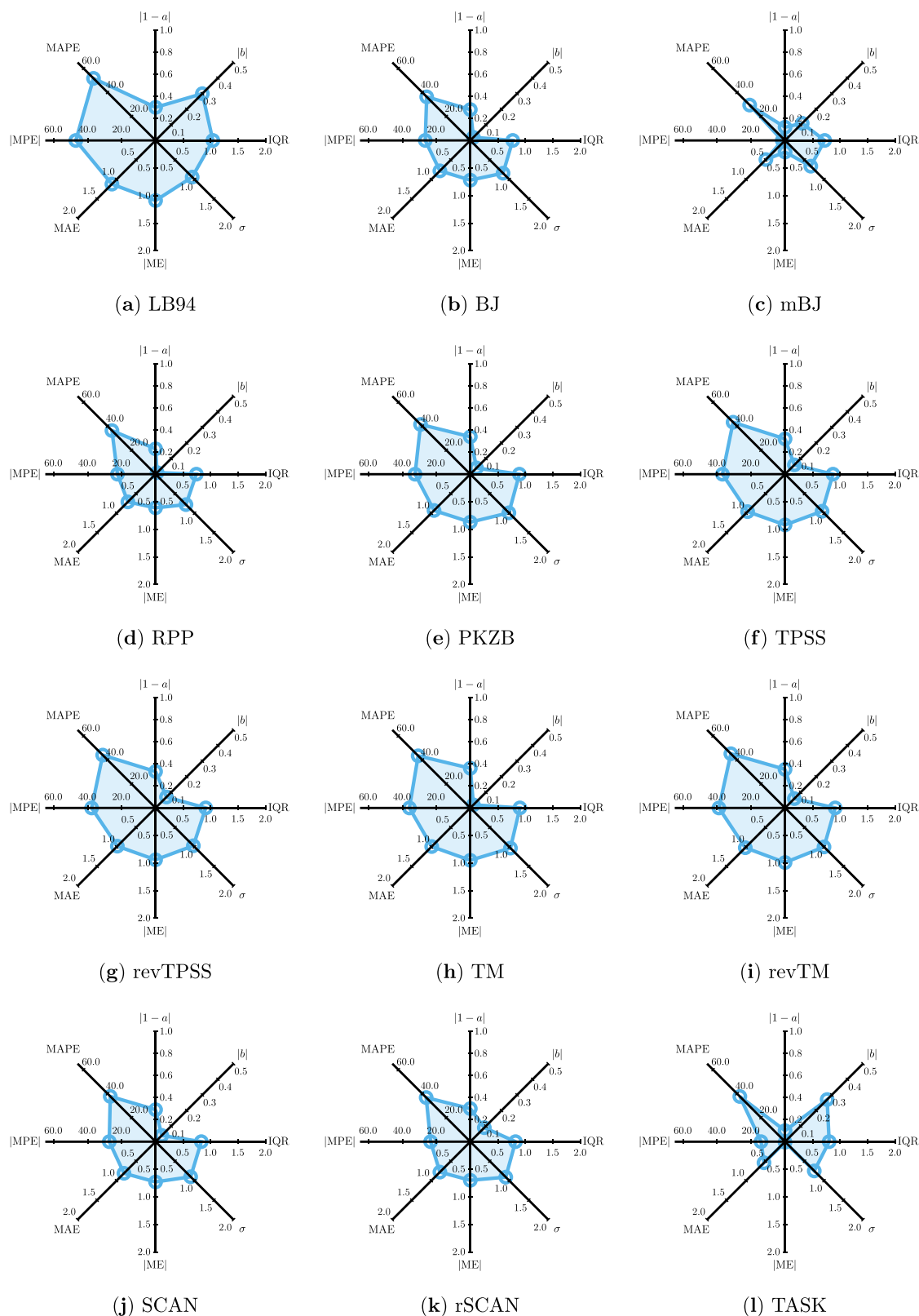


Fig. 2 Radar plots displaying the statistical quantities calculated in this benchmark (MAE, ME, $|b|$, σ , and IQR in eV; MPE and MAPE in %). Here we show the radar plots of LB94 and a set of meta-GGA functionals.

the range of small to intermediate band gaps (<4 eV). As in the case of the previous GGAs, the mean errors in the fd subset are smaller than those in the sp one, although the situation is inverted for the percentage errors. As seen in Supplementary Fig. S7, this is

due to the prevalence of smaller gaps in the fd subset. With 17 false metals predicted, it performs very close to Slocc on this metric.

We end the discussion of the results at the GGA rung with AK13LDA (Fig. 1i). With M(A)PE of 14% (38%) and M(A)E of 0.05 eV

(0.58 eV), this is the most accurate GGA among the functionals tested in this work. Its very small ME is an indication of the centered nature of the error distribution. As it turns out, this result is not achieved through a generally good behavior, but this is rather due to error cancellation between the different subsets. For *sp* materials, AK13LDA tends to overestimate band gaps, particularly the small ones below 2 eV. For *fd* materials, a general underestimation is observed. As a result, the overall distribution of percentage errors is almost symmetrical (Supplementary Fig. S7). With only 10 false metals predicted, it is the most reliable GGA in identifying semiconductors. All in all, this makes AK13LDA the best GGA functional tested in this work, in particular for band gaps larger than 1 eV. Its performance is overall rather similar (albeit slightly worse for the M(A)PE) to another GGA that we tested in ref. ⁴³, the HLE16. It should be noted that, depending on the computer code, calculations with AK13LDA can be subject to numerical difficulties, particularly in cases of large gradients, and the self-consistent field convergence may be hard to reach. This type of convergence behavior is not uncommon with functionals that exhibit bumpy or unusual potentials, like AK13LDA (see plots of the AK13LDA potential in ref. ¹³⁴).

All considered GGAs, except AK13LDA, present an overall tendency to underestimate band gaps. Considering chemical compositions, GGAs exhibit their largest errors for compounds containing Ni, Pd, Pt, and Pb (see Supplementary Fig. S1). Curiously, AK13LDA seems to be the functional suffering the most from this condition, in spite of its overall better performance.

meta-GGA

Turning now to the meta-GGAs, we start with PKZB (Fig. 2e). This functional presents mean errors very similar to SOGGA11, nevertheless PKZB reduces the number of predicted false metals (from 28 to 20), particularly in the *fd* subset. Also in the same way as SOGGA11, PKZB performs poorly for Ni and Pt compounds (Supplementary Fig. S1).

TPSS, revTPSS, TM, and revTM (Fig. 2 f–i), as it turns out, present very similar results to PKZB. This is not only inferred from their mean errors, but also from a direct comparison of the calculated results (Supplementary Fig. S8). We just note that all of these functionals predict a larger number of false metals than PKZB (20 for PKZB, 28 for TPSS, 30 for revTPSS, 24 for TM, and 30 for revTM), and that revTPSS and revTM both predict more false metals than their respective parent functionals TPSS and TM.

MVS (Fig. 3a) stands out as one of the few functionals tested here with positive MPE (2.4%). This is due to the large overestimation of band gaps smaller than 1 eV, while beyond this range it underestimates band gaps, but performs better than the previous meta-GGAs. It also performs clearly better for the prediction of false metals, 11 in total.

As expected, rSCAN (Fig. 2k) performs overall similarly to SCAN⁴³, rSCAN achieves a M(A)PE of -23% (37%) while SCAN leads to -27% (38%). A noticeable difference is the description of small band gaps. For band gaps smaller than 1 eV, the MPE of SCAN is about 25%, while it is reduced to 10% for rSCAN. For both functionals, the MAPE for these small band gaps is large (75%). From an elemental point-of-view, the largest difference is that rSCAN presents much larger errors for Pt compounds (with a MAPE of 180%) than SCAN (MAPE of 51%). rSCAN improves the number of predicted false metals, 15, from the 20 of SCAN.

Among the newly tested functionals, HLE17 is one of the best performers (Fig. 3g). In spite of being a very simple (empirical) rescaling of TPSS, it presents considerably better results than all meta-GGAs discussed so far (PKZB, TPSS, etc.) in every measure. Not only does HLE17 reduce the underestimation of the band gaps (ME of -0.44 eV), it does so consistently across the *sp* and *fd* subsets (ME is -0.42 and -0.45 eV, respectively). On the other hand, such good results from HLE17 are not really unexpected,

since, as already mentioned above, this functional was empirically fitted to be accurate for band gaps. Supplementary Figure S11 shows the comparison of the calculated gaps between TPSS and HLE17. Although a linear trend is observed, we can witness a larger dispersion of the results than in the previous cases. A linear fit to the (TPSS/HLE17) data yields $a = 1.1$ and $b = 0.3$, from which we clearly see the HLE17 tendency to predict larger band gaps than TPSS. Both of these observations were expected, as HLE17 increases the exchange contribution with respect to TPSS, leading to increased band gaps. Apart from this, other measures such as the standard deviation σ (0.82 eV), the IQR (0.75 eV), the MnE (-0.30 eV), and the MAE (0.60 eV) are also improved throughout the different subsets. Most striking are the changes in the mean (absolute) percentage errors which go from -37% (43%) for TPSS to -11% (31%) for HLE17. These results come mostly from the improved description of small band gap *sp* materials (Supplementary Fig. S12), which also explains the tiny MPE of 0.28% in this subset. Overall, the good quality of HLE17 is also visible when comparing the TB67 subset among the functionals. With 17 predicted false metals, HLE17 performs at the same level as Sloc in this aspect.

From Fig. 2n, the TASK functional seems very close to AK13LDA in terms of performance. This is confirmed by direct comparison in Supplementary Figs. S14 and S15, which in turn makes the overall discussion of this functional very similar to that of AK13LDA. However, these functionals present a different behavior in the TB67 subset. Here, TASK performs similarly to PBE0, overestimating small gaps and underestimating large ones, but with a much better performance for mid-range gaps. This can be understood by looking at the errors in this subset: M(A)E -0.4 eV (0.8 eV) and M(A)PE 14% (41%). Apart from this, TASK and AK13LDA present a similar elemental error fingerprint (with Ni, Pt and Pb showing the largest errors) and number of false metals (11 and 10, respectively).

Considering now the meta-GGAs of the Minnesota family, M06-L (Fig. 3b) shows a fingerprint very similar to PKZB, TPSS, revTPSS, TM, and revTM, but with improved results. Actually, with M(A)E of -0.8 eV (0.8 eV) and M(A)PE of -30% (38%), M06-L shows results rather similar to SCAN and the BJ potential, which were tested in ref. ⁴³ (the results are reproduced in Fig. 2l and b, respectively).

The functional revM06-L (Fig. 3c) performs very differently from its parent M06-L, but much more akin to MVS, as shown in Supplementary Fig. S16. However, comparing the M(A)E and M(A)PE of revM06-L and HLE17, the results are of similar quality, and actually the number of false metals is only 10 for revM06-L (17 for HLE17). Thus, as HLE17, revM06-L should be considered as a functional with improved accuracy with respect to most other meta-GGAs. The largest errors with revM06-L are for compounds with Ni and Pt atoms (Supplementary Fig. S1).

The MA(P)E of M11-L are somehow similar to M06-L, i.e., worse than HLE17 and revM06-L but better than TPSS or revTPSS, for instance. Noteworthy, M11-L presents the same overestimation as MVS and revM06-L for band gaps smaller than 1 eV, which is responsible for the disparity between MAPE (40%) and MPE (-4%).

Finally, we turn to MN12-L and MN15-L (Fig. 3e, f), which perform similarly, although the latter is slightly better than the former for the M(A)E and MP(A)E. This can also be seen by comparing directly the results of the two functionals (Supplementary Fig. S18). These two functionals are overall inferior to HLE17.

We finish with the RPP potential (Fig. 2d) that, with M(A)E of -0.6 eV (0.7 eV) and M(A)PE of -22% (36%), performs pretty much as rSCAN (and therefore also SCAN).

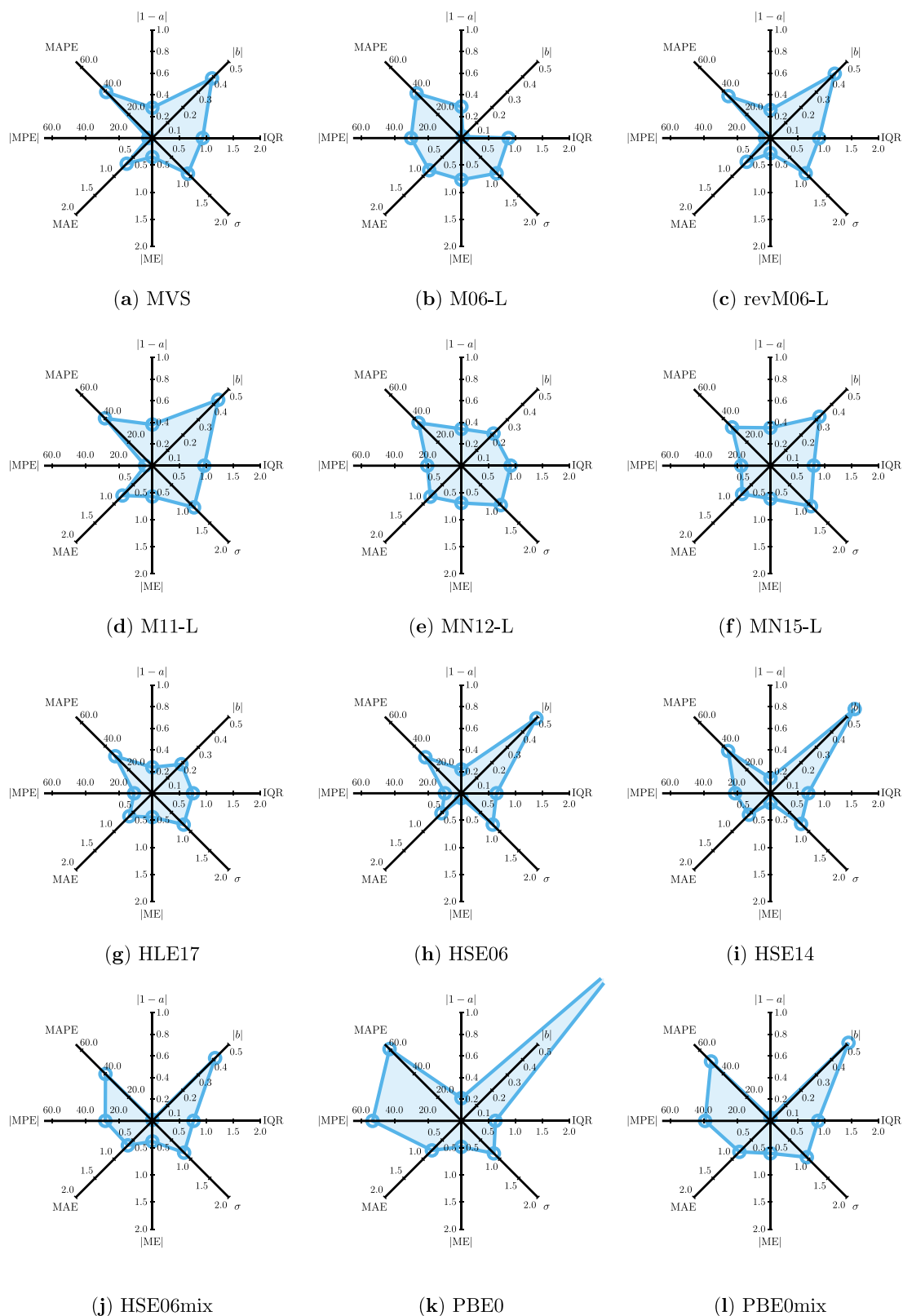


Fig. 3 Radar plots displaying the statistical quantities calculated in this benchmark (MAE, ME, $|b|$, σ , and IQR in eV; MPE and MAPE in %). Here we show the radar plots of a set of meta-GGA and hybrid functionals.

Summary for the MAE and MAPE

The MAE and MAPE of most functionals tested in the present and previous⁴³ works are shown in Fig. 4. This visual summary clearly

shows which functionals reproduce most accurately the experimental results. The mBJ potential and the hybrid HSE06 are the most accurate for both the MAE and MAPE. Looking at the MAPE,

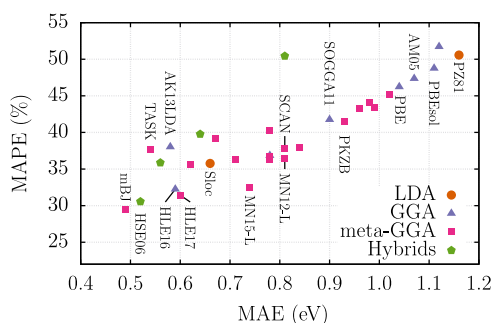


Fig. 4 Overview of the MAE and MAPE of all studied functionals. MAE (x axis) vs. MAPE (y axis) for the functionals included in this work and in ref. ⁴³.

the five best performing functionals are mBJ, HSE06, HLE17, HLE16, and MN15-L, while considering the MAE, we can pick out mBJ, HSE06, TASK, HLE14, and AK13LDA. Considering its extremely simple mathematical form, the performance of Sloc is very interesting. For comparison, we should mention that the results from ref. ⁷⁷, obtained on the same data set with the computationally fast GLLB-SC potential¹³⁵, are 0.7 eV and 39% for the MAE and MAPE, respectively, which is very similar to MVS or EV93PW91, for instance. Also, among the five hybrid functionals tested in ref. ⁴³, HSE06 is the most accurate. At that point, we also remind that the mBJ potential shows also good performance (similar to HSE06) for antiferromagnetic oxides with localized 3d-states (not considered in the present work), while all other GGAs or meta-GGAs are clearly worse for these systems^{16,17,76}.

Other metrics

Using other metrics may provide another way of comparing the performance of the functionals or reveal trends. For instance, the standard deviation σ lies in the range from 0.75 eV (TASK) to 1.09 eV (M11-L). This indicates that TASK, but also RPP, lead more systematically to the same error than the other functionals. A similar dispersion is observed for the IQR, which varies from ~ 0.75 eV (RPP and HLE17) to 1.04 eV (LB94), while the MADM ranges from ~ 0.36 eV (RPP and HLE17) to 0.52 (LB94). The MnE presents some more interesting details as AK13LDA is the only functional (along with all hybrid functionals tested in ref. ⁴³) with a positive (and very small) value for this quantity (0.07 eV). This is not surprising given the overestimation of band gaps smaller than 1 eV, in combination with the small errors for larger band gaps, as previously discussed. From the remaining functionals, including those tested in ref. ⁴³, HSE06 (0.0 eV), revM06-L (−0.1 eV), and mBJ (−0.1 eV) present the smallest absolute values of MnE.

From the point of view of reproducing the experimental trends with a linear fit ($y = ax + b$), the functionals that lead to the smallest error in this work are Sloc, AK13LDA, and TASK. Sloc is rather well balanced, since with coefficients $a = 0.82$ and $b = 0.04$ eV, it is simultaneously near the optimal values for both quantities. AK13LDA and TASK on the other hand yield values of a of 0.96 and 0.90, respectively, whereas most other functionals give values of a in the range 0.62–0.77. Their intercepts ($b = 0.15$ and 0.27 eV, respectively) are however worse than the one of Sloc. Since b is typically associated with systematic errors, AK13LDA could be easier to improve than Sloc. Among the methods tested in our previous work⁴³, the mBJ potential shows the best linear fit, with $a = 0.88$ and $b = 0.10$ eV, which is overall of the same quality as Sloc and AK13LDA.

In the *sp* and *fd* subsets of compounds, the Sloc and AK13LDA functionals have similar behaviors. For *sp*, they maintain similar values of a (0.84 for Sloc and 0.97 for AK13LDA), but worse values for b (0.20 for Sloc and 0.40 for AK13LDA). However, for the *fd* subset, both functionals perform significantly worse.

Functional reparametrization

In the “Results and discussion” section and in ref. ⁴³, we presented a large quantity of statistical data regarding the calculation of band gaps using different functionals. Now we instead study systematically the performance of a few selected functionals for band gaps as a function of the internal parameters of the functional. This allows not only to optimize the functionals, but also to better understand their overall behavior. For this purpose, we selected at least one functional from each rung, with the exception of the hybrids, due to their computational cost.

Due to the large number of calculations required for this task, we did not use the complete data set. Instead, we decided to construct a subset comprising the 85 materials (listed in Supplementary Table SII) with the smallest number of electrons and that do not exhibit strong spin–orbit coupling. Although essential for computational reasons, we are well aware that this choice may lead to a bias. We note nevertheless that a large number of tests and benchmarks are often conducted with similar small data sets^{76,86,136–138}.

Using this reduced data set, we selected the best parameters to optimize the size of the band gap. The functionals defined with these fitted parameters will be denoted with the prefix “t-”, to distinguish them from their regular counterparts. Radar plots depicting the errors of the fitted functionals can be found in Fig. 5.

We begin our analysis with Sloc, given its surprisingly good results and its extremely simple form. In Fig. 6, we present the MAPE as a function of the two parameters of Sloc, namely a and b (Eq. 4). In the figure, we also show the parameters that correspond to the standard LDA exchange functional and to the original Sloc. It turns out that the MAPE depends much more strongly on the prefactor a than the exponent b . As Sloc underestimates band gaps, one expects the optimal parameters to counterbalance this by increasing the weight of the exchange contribution. This is achieved by increasing a and decreasing b . Indeed, the minimum of the MAPE is located at a rather small region of the space of parameters, specifically at $a = 1.775$ and $b = 0.260$. These new parameters define t-Sloc.

The fitted functional (Fig. 5a and Supplementary Fig. S23) indeed reduces the underestimation of Sloc but only by 0.05 eV on average for the complete data set (individual corrections vary from −0.18 eV for AgF to 0.80 eV for Ar). With the exception of the ME and MPE, the other metrics, most notably the MAPE, are however not improved upon.

The PBE approximation is certainly one of the most successful and used functionals in the solid-state physics community. It depends on four parameters whose numerical values are fixed by theoretical considerations, specifically (i) μ (to obey the slow varying limit of exchange), (ii) κ (to enforce the local Lieb–Oxford bound¹³⁹), (iii) β (to obey the slow varying limit of correlation), and (iv) γ (to cancel the logarithmic singularity of the correlation in certain limits). The flexibility offered by this parametrization has originated a series of other members of this family (e.g., RPBE⁸², revPBE¹⁴⁰, APBE¹⁴¹, PBEsol⁶⁰, PBEfe¹⁴, xPBE¹⁴², and others¹⁴³), some of which fitted to give more accurate results for specific quantities.

The two parameters that are the most relevant for band gaps are the ones that define the exchange contribution, specifically μ and κ . In Fig. 7, we show the MAPE as a function of these two parameters. The values of μ and κ for a few members of the PBE family of functionals are also indicated in the picture. The MAPE surface is considerably more complex than with Sloc, a feature also observed for other quantities¹⁴⁴. There is a clear valley at around $(\mu, \kappa) = (0.6, 3)$ and some craters at higher κ .

We chose to study further two different minima present in the plot, namely the values of $(\mu_1, \kappa_1) = (2/3, 3 + 1/6)$ (denoted t-PBE1) and $(\mu_2, \kappa_2) = (1/3, 12 + 5/6)$ (denoted t-PBE2, not shown in the

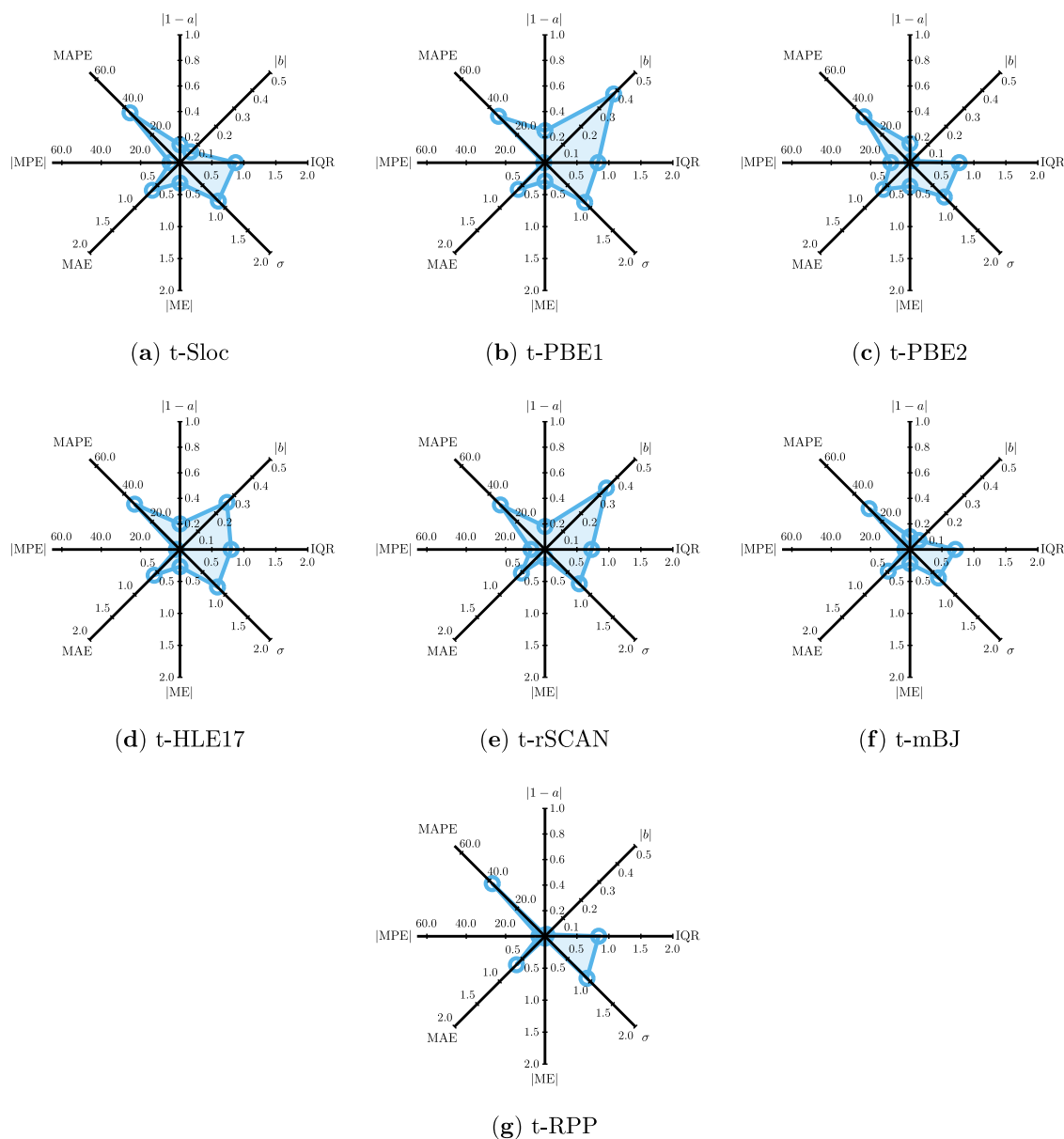


Fig. 5 Radar plots displaying the statistical quantities calculated in this benchmark (MAE, ME, $|b|$, σ , and IQR in eV; MPE and MAPE in %). Results for the reoptimized functionals: (a) t-Sloc, (b) t-PBE1, (c) t-PBE2, (d) t-HLE17, (e) t-rSCAN, (f) t-mBJ, (g) t-RPP.

figure). Note that these exhibit considerably higher values of κ with respect to typical PBE reparametrizations.

When compared to PBE, both parametrizations improve upon the severe band gap underestimation and the error dispersion (Fig. 5b, c), and reduce significantly the number of false metals (31 for PBE, 10 for t-PBE1, and 12 for t-PBE2). These two fitted functionals behave very differently for different band gap ranges. For band gaps smaller than 5 eV, t-PBE1 predicts larger band gaps than t-PBE2 (Supplementary Fig. S26), being closer to experimental values in this range. For the higher range this situation seems to be inverted, with t-PBE2 providing a better description of the gaps. However, the reduced number of compounds in this region makes it difficult to ascertain whether this is an actual improvement or only specifically for these compounds.

We now turn to the meta-GGA HLE17. Being simply a rescaling of the TPSS energy functional (Eq. 11), we study its behavior by allowing arbitrary weights w_x and w_c for each of the components,

$$E_{xc}^{\text{t-HLE17}} = w_x E_x^{\text{TPSS}} + w_c E_c^{\text{TPSS}}, \quad (14)$$

to be optimized. In addition to HLE17, we apply the same procedure to rSCAN, i.e.,

$$E_{xc}^{\text{t-rSCAN}} = w_x E_x^{\text{rSCAN}} + w_c E_c^{\text{rSCAN}}. \quad (15)$$

Figure 8 represents the MAPE calculated on the training set, as a function of the fitting parameters for both scaled HLE17 and rSCAN. For both functionals the variation of MAPE is smooth, although there are qualitative differences between them. For HLE17, the MAPE is essentially independent of the correlation term. For fixed w_c , the MAPE has an asymmetric shape, increasing more rapidly for smaller w_x . In the end, we selected $w_x = 1.35$ and $w_c = 1.00$ for the final version of t-HLE17.

In the case of rSCAN, correlation is much more relevant than in t-HLE17. This can be seen from the well defined minimum in the MAPE surface, located approximately at $w_x = 1.30$ and $w_c = 1.40$. These values define the t-rSCAN functional.

Regarding t-HLE17 (Fig. 5d), the small variation in the weight of the exchange part leads to slightly increased band gaps with respect to HLE17 (Supplementary Fig. S20). This is also visible from

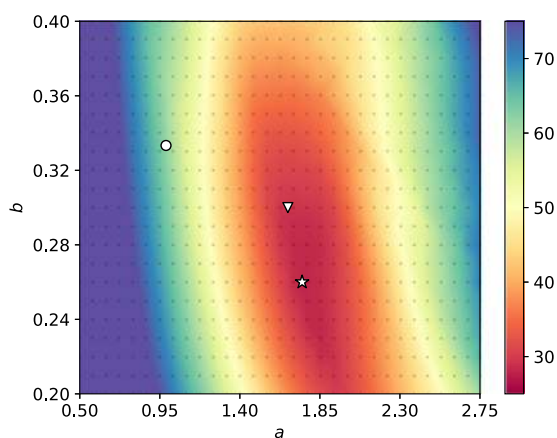


Fig. 6 MAPE as a function of the parameters a and b of Sloc. The symbols indicate the values of the parameter for the different functionals of this family: ∇ Sloc, \circ LDA, and \star t-Sloc. The color-coded scale on the right is in %.

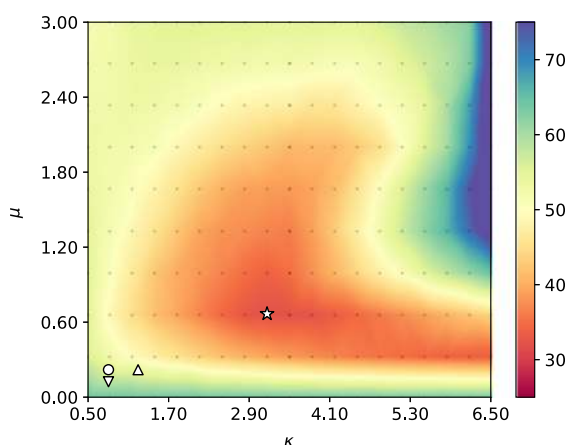
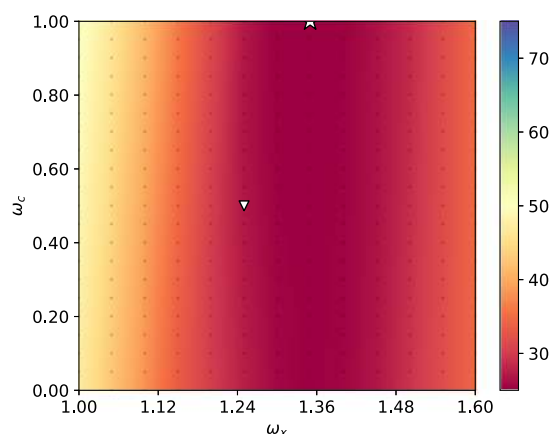


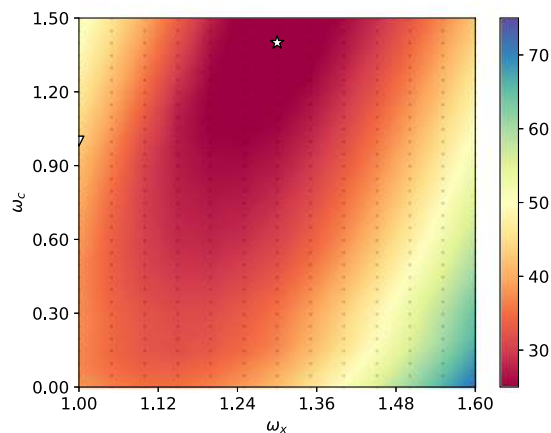
Fig. 7 MAPE as a function of the parameters μ and κ of PBE. The symbols indicate the values of the parameter for the different functionals of this family: PBE (\circ), revPBE (Δ), PBEfe (∇), and t-PBE1 (\star). The color-coded scale on the right is in %.

the change in the ME for the whole data set (from -0.44 eV for HLE17 to -0.27 eV for t-HLE17). However, such increase is not universal, and the amount of compounds for which t-HLE17 predicts a smaller band gap than HLE17 is also not very small (63 compounds in total, with LaF_3 showing the largest reduction by 1 eV). As usual, the increase in predicted band gaps leads to an overestimation of band gaps smaller than 1 eV, which in turn translates to a large increase in the MPE. This increase is clearly visible in the sp subset which goes from 0.28% in HLE17 to 10% in t-HLE17. Values of absolute errors (MAE and MAPE) are not significantly changed between the two functionals. Finally, all of these changes are not accompanied by a change in error dispersion, as t-HLE17 presents a σ of 0.83 eV and IQR of 0.80 eV, very close to the HLE17 values.

The improvements of t-rSCAN over rSCAN are much more significant than the corresponding ones for HLE17. This fact is easily seen in Supplementary Fig. S21. As expected, t-rSCAN reduces the underestimations of rSCAN, and does it in all subsets, bringing the ME to -0.13 eV (*all*), -0.09 eV (sp) and -0.17 eV (fd). Unlike the case of t-HLE17, the MA(P)E are also significantly improved. The only error metric where t-rSCAN worsens with respect to rSCAN is in the MAPE of the sp subset (35% vs. 33%, respectively). When looking at the distribution of errors along the band gap range, one understands that this is, once again, due to



(a) t-HLE17



(b) t-rSCAN

Fig. 8 Color map of the MAPE as a function of functional parameters of HLE17 and rSCAN. MAPE as a function of the parameters ω_x and ω_c of HLE17 (a) and rSCAN (b). We also indicate the actual parametrization of the original (∇) and fitted (\star) versions of the functionals. The color-coded scale on the right is in %.

the overestimation of band gaps smaller than 1 eV. Apart from this, there are some small gains in error dispersion and a reduced number of false metals (from 17 to 9). The description of the a and b coefficients is also not particularly good for both rSCAN ($a = 0.70$ and $b = 0.09$) and t-rSCAN ($a = 0.82$ and $b = 0.34$). In addition, this functional maintains the same problems as rSCAN for compounds containing Ni, Pt, Pd, and Pb (Supplementary Fig. S1).

The mBJ potential is an exchange- and potential-only meta-GGA approximation, which describes v_x as an enhancement of the BJ potential⁶¹,

$$v_x^{\text{mBJ}} = cv_x^{\text{RR}} + (3c - 2) \frac{1}{\pi} \sqrt{\frac{5}{12}} \sqrt{\frac{2\tau}{n}}, \quad (16)$$

with

$$c = a + \beta \left[\frac{1}{V_{\text{cell}}} \int_{\text{cell}} \frac{|\nabla n(\mathbf{r})|}{n(\mathbf{r})} d\mathbf{r} \right]^{1/2}. \quad (17)$$

The two parameters $a = -0.012$ and $\beta = 1.023 \text{ Bohr}^{1/2}$ were originally obtained by fitting band gaps of all-electron calculations to experimental ones¹⁵.

The MAPE dependence on a and β is shown in Fig. 9, where we can observe a long diagonal band where the MAPE is essentially constant. The original parameters are already located at the center

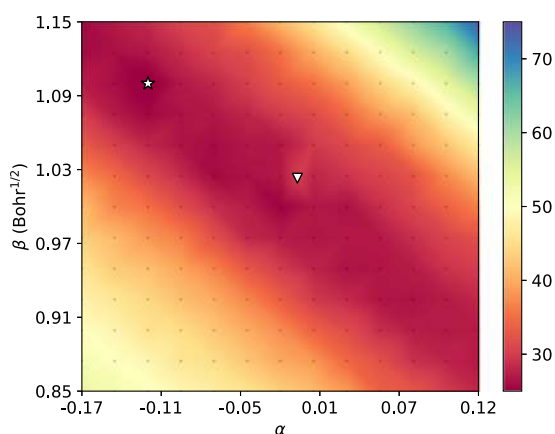


Fig. 9 Color map of the MAPE as a function of the parameters α and β of mBJ. The scale on the right is given in %. The symbols indicate the actual parameter values for the different functionals of this family: mBJ (∇) and t-mBJ (\star).

of this valley. The absolute minimum of the MAPE is $\alpha^{\text{t-mBJ}} = -0.125$ and $\beta^{\text{t-mBJ}} = 1.100$ Bohr^{1/2}, which we took to define t-mBJ. As we can see, the β parameter is very close to the original one of mBJ, while α is more negative. This, however, does not seem to have a significant impact on the predicted band gaps, as the metrics of mBJ and t-mBJ are essentially the same, as can be seen by comparing Figs. 2l and 5f. The most notable difference between mBJ and t-mBJ is found for Ne (experimental gap of 21.48 eV) for which mBJ predicts 18.85 eV, whereas t-mBJ predicts 21.35 eV. However, it is worth mentioning that for the very large band gaps above ~ 12 eV, mBJ results obtained with VASP are clearly smaller (up to a few eV for Ne) than the reference mBJ values obtained with an all-electron code⁷⁶. Such large underestimations by VASP is specific to mBJ and is, according to our investigations, mainly due to the calculation of c [Eq. 17].

As previously mentioned, the RPP potential was developed to solve some of the problems of the BJ potential. Therefore, it is rather tempting to modify the original form of RPP by following the ideas of mBJ in order to describe more accurately the band gaps (see also ref. ¹³⁴):

$$v_x^{\text{t-RPP}} = cv_x^{\text{BR}} + (3c - 2) \frac{1}{\pi} \sqrt{\frac{5}{12}} \sqrt{\frac{D}{n}}, \quad (18)$$

where c is given by Eq. (17) as in mBJ. Overall, we observe in Fig. 10 that the resulting MAPE looks very similar to the one for mBJ in Fig. 9, but we remark that a different scale is used for the parameters α and β .

The new potential t-RPP shows very promising results. Looking at Fig. 5g, we see that it presents excellent values for the a and b coefficients of the linear fit, at 0.99 and 0.01 eV, respectively. The results are slightly worse for b in the sp subset ($a = 1.00$, $b = 0.13$ eV), but more clearly for a and b in the fd subset ($a = 0.81$, $b = 0.23$ eV). The mean errors are also very small (ME of -0.03 eV and MPE of 3.1%), but the absolute mean errors (MAE of 0.63 eV and MAPE of 38%) are comparable to other functionals like AK13LDA or MVS. The reason for this comes from the dispersion of errors, and the fact that t-RPP both underestimates and overestimates band gaps, depending on the band gap range under consideration.

Summary for the fitted functionals

A summary of the MAE and MAPE for all fitted functionals is presented in Fig. 11, where the original functionals are also shown for comparison. The most impressive improvements are obtained by reoptimizing the PBE functional, so that the performances of t-PBE1 and t-PBE2 are similar to HLE16 and (t-)HLE17. Noteworthy,

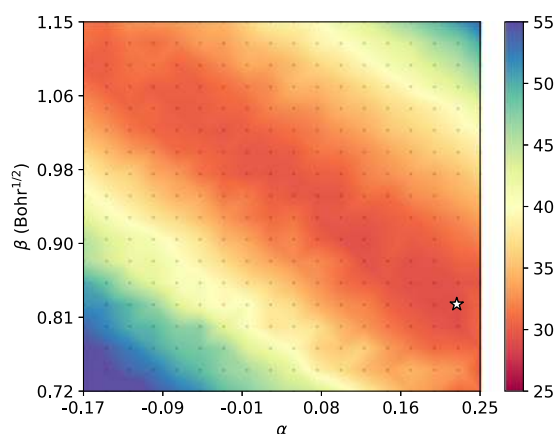


Fig. 10 Color map of the MAPE as a function of the parameters α and β of t-RPP. The scale on the right is given in %. Note the different scale with respect to the one for (t-)mBJ in Fig. 9.

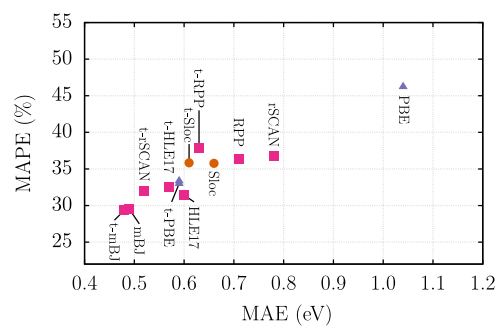


Fig. 11 MAE (x axis) and MAPE (y axis) on the band gap for the fitted functionals in this work. The errors of the original functionals are also shown for comparison.

the fitted PBE functionals still satisfy the homogeneous electron gas limit, while HLE16, (t-)HLE17, and also t-rSCAN do not, since they have scaled exchange and correlations parts.

The improvement of rSCAN is also significant. As somehow expected, for functionals which originally were already fitted specifically for band gaps (mBJ, Sloc, and HLE17), the improvement is only minor. We expect that the same conclusion may probably hold also for HLE16.

Machine learning

We witnessed in the previous section the difficulty in improving existing functionals for band gaps, and in particular the persistence of the MA(P)E in the large data set. We will now take one step further and build more complex machine-learning models that combine the results of multiple functionals to achieve higher accuracy.

We chose two different methods that provide highly interpretable models, as the small size of the data sets severely limits the complexity of the models. Specifically: (i) The sure independence screening and sparsifying operator (SISSO)¹⁴⁵ starts with a number of features, combines them with linear and non-linear operators to create millions of higher-level features and then attempts to find the best sparse linear model using these higher-level features. (ii) The model agnostic supervised local explanations (called MAPLE)¹⁴⁶ is a recently developed machine-learning technique that combines the usability and accuracy of decision trees with the interpretability of local linear models and self-explanations, therefore circumventing the usual accuracy-interpretability trade-off. MAPLE uses traditional decision trees,

e.g. random forests¹⁴⁷, to arrive at a local simulatable¹⁴⁸ linear model. Furthermore, it can output the most influential training samples for a prediction as a form of self-explanation.

As starting features, we used all the results calculated with the various functionals from the benchmark data set (the fitted functionals were not included). Furthermore, we complemented this data with composition specific features created with *matminer*¹⁴⁹ (see Supplementary Table SIII for a full list), with the symmetry groups in Hermann–Mauguin notation, as well as the crystal structure prototype according to the ICSD¹⁵⁰. In the case of MAPLE, the feature importance of the underlying random forest model can be used to decrease the number of features by always removing the least important functional and refitting the model. While SISSO ideally selects its features completely independently, the feature selection task in this case is highly non-trivial as all the functionals are highly correlated. To keep the number of functionals that end up in the final features, and therefore the computational cost, limited, we used the most important functionals selected by MAPLE as a starting point for SISSO. Unsurprisingly, the mBJ functional was consistently chosen as the most important or second most important functional. Other relevant functionals varied, for example, we obtained the HLE16 and RPP/M06-L or the LDA and LDA-SOC (with spin-orbit corrections). Due to the high variability in the systems and their band gaps and the small amount of data, 10-fold cross-validation was used to evaluate the models instead of a separate test set. The training/validation split was chosen as 70%/30%. All the following results are averaged according to this validation scheme.

SISSO only returned results comparable to MAPLE once higher-dimensional models were used, therefore we will concentrate on the MAPLE models in the following. For combinations of two functionals and three functionals together with the selected experimental features, we selected the best performing pair of functionals by brute-force trying all combinations of functionals together with mBJ. In Fig. 12, we compare MAPLE models using various numbers of functionals and elemental properties as input to the best functional in the complete benchmark (present work and ref. 43). Using only band gaps calculated with the mBJ functional as input, the MAE, IQR, and ME are significantly improved while the MAPE and MPE increase. Using band gaps calculated with two or three functionals the MAE, ME, standard deviation σ , and IQR are significantly improved, while the MAPE still remains nearly unchanged. Concerning the linear fit, the coefficient b is also clearly improved, while a is increased. Examining the distribution of the MAPE reveals that this quantity is larger than 70% for the small band gaps (<1 eV) while being comparatively small at ~20% for the larger band gaps. We also attempted to learn the error of the mBJ functional in comparison to the experimental data (which is known as delta-learning, or using a crude estimator of property¹⁵¹). However, in our case this approach did not seem to work.

The difficulty in improving the estimators, either by reoptimizing functionals or through machine learning models, is perhaps not surprising if we consider the relatively small size of the data set (473 materials), the large range of band gaps (between 0 and 20 eV), and the large diversity of compositions (nearly all chemical elements) and crystal structures (over 150 structure prototypes, from elementary to quaternary materials). However, if we only consider the region between 0.9 and 2 eV (171 entries in total), which is the most relevant for applications like photovoltaics, we can actually achieve a significant improvement in all statistical properties. This can be clearly seen from Fig. 13, where we depict the radar plots for three models, one including compositional features together with the mBJ band gaps, one adding the RPP band gap and another one adding the HLE16 and rSCAN band gaps. The MAE and MAPE are both improved by 40% in comparison to the mBJ functional and present the most accurate estimator of band gaps between 0.9 and 2 eV that we could find.

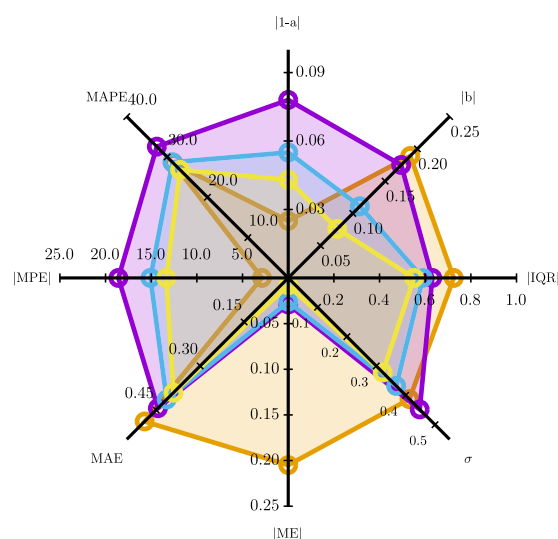


Fig. 12 Radar plot for three MAPLE models and the mBJ functional. All models use various elemental features: purple uses the mBJ band gap, light blue uses mBJ and M06-L band gaps, and yellow uses mBJ, HLE16 and M06-L band gaps. Results correspond to 10-fold cross-validation with 70%/30% training/test split. Note the different scale with respect to the previous radar plots. The radar plot of the MBJ functional is given in orange for comparison.

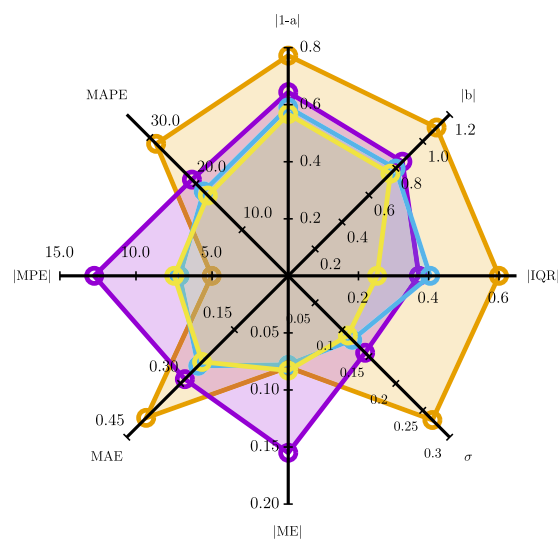


Fig. 13 Radar plot for three MAPLE models and the mBJ functional. All models use various elemental features: purple uses the mBJ band gap, light blue uses mBJ and RPP band gaps, and yellow uses mBJ, HLE16 and rSCAN band gaps. The radar plot of the MBJ functional is given in orange for comparison. The results are for only the materials with band gaps from the region of interest between 0.9 and 2 eV. Results correspond to 10-fold cross-validation with 70%/30% training/test split. Note the different scale with respect to the previous radar plots.

In order to allow other researchers to improve their band gap predictions, we will publish the machine learning model on our website (<https://www.tddft.org/bmg/>). Besides predicting the band gap, the MAPLE model is able to highlight the materials in the training set that were most relevant for its prediction as well as the local linear fit from which the band gap is calculated. This allows researchers to make an informed decision on whether to trust the predicted band gap.

In summary, we performed calculations of the electronic band gap of 473 materials using 21 different approximations to the exchange-correlation functional of density-functional theory. These include a series of local, generalized-gradient, and meta-generalized-gradient approximations. Together with our previous results of ref. ⁴³, this amounts to the largest benchmark of band gaps up to date, and allows us an unprecedented view on the problems and deficiencies of density-functional theory to calculate this important physical property.

Statistically, at the lowest rung of Jacob's ladder of density functionals, the most accurate approximation is Sloc. Going up the ladder to the second and third rungs, we then encounter the GGA HLE16 and the meta-GGAs mBJ, HLE17, and TASK. Finally, the hybrid HSE06 (fourth rung) is also among the most accurate. Overall, the most reliable functionals for band gap calculations are mBJ, that yields the lowest mean absolute (percent) error, and HSE06. Other semilocal functionals that perform reasonably well are for instance AK13LDA and some meta-GGAs of the Minnesota family like revM06-L.

Even if our objective was not to perform a detailed analysis of the computational cost of the benchmarked functionals, as this would require to carefully control the running conditions for all calculations, we believe that some information concerning the runtimes can be useful to the readers. Our computational resources consisted in a shared cluster. To make the best possible use of the available resources, not all calculations could be performed with the same number of cores, type of processors and parallelization options. Furthermore, our running times refer to VASP and could be substantially different if other implementations are used. For example, the adaptively compressed exchange operator algorithm¹⁵² provides a significant boost of hybrid functional calculations, but it is not implemented in version 5.4.4 of VASP.

Keeping this in mind, we can compare the numerical efficiency of calculations using different functionals for our data set. The average meta-GGA is around 5 times slower than LDA/GGAs, probably due to the need to evaluate extra terms in the Hamiltonian. The mBJ (and related functionals) are 10–50 times slower, due to the larger number of iterations required for convergence and to the time needed to solve the non-linear equation that appears in the definition of the functional. Calculations using regular hybrids and dielectric-dependent hybrids from ref. ³⁹ are two orders of magnitude slower than LDA/GGAs calculations.

We also analyzed a smaller subset of our 473 materials that includes many of the materials that are commonly used in the benchmarks of band gaps. It includes mostly elementary or binary zincblende and wurtzite semiconductors, with a strong presence of *sp* elements. The small set is mostly composed of “simple” materials and is therefore not fully representative of the set of semiconductors and insulators. Fitting functionals to this set may therefore lead to a bias.

In order to understand the behavior of the analytical form of selected functionals, we then investigated how the average error in the band gap depends on the individual parameters used in the construction of the functionals. We performed such maps for Sloc, PBE, HLE17, rSCAN, mBJ, and RPP. By using the values of the parameters that minimized the mean absolute error, we then constructed reparametrizations of these approximations. It turns out that one can construct accurate functionals for band gaps at all rungs of Jacob's ladder if the fraction of exchange is increased. Of course, this may lead to a degradation of the results for other physical quantities (see ref. ⁷⁷ for a discussion on that problem). Although one can improve considerably some of the statistical quantities describing the error in the band gaps, the mean average percent error remains high, likely due to the small size of the fitting set.

Finally, we developed machine learning models to improve the prediction of band gaps. The input for our models are band gaps

calculated with approximate exchange-correlation functionals and structural and compositional data. We were unable to train consistently better models that worked in the whole range of band gaps. This can be explained by the small size of the data set (for machine learning applications) and by the extremely large diversity of the materials considered. We could, however, train a very simple model that improves considerably the band gaps in the range of 0.9–2 eV, an interval that includes the range relevant for photovoltaic applications.

The recurring problem in benchmarking band gaps is the relatively small number of reliable experimental data. These can be complemented by information on new materials, but also by revisiting “old” semiconductors with more modern experimental techniques. This can lead to a deeper understanding of material properties and to the development of better theoretical tools and approaches. For band gaps, these obtained data should also ideally be corrected for temperature effects, zero-point corrections, excitonic binding energies, etc., factors that together can explain some of the deviation to calculated results. This time-consuming and tedious process can nevertheless be accelerated by the use of active-learning techniques to propose new candidates for experimental characterization to better cover material space. A close collaboration between experimentalists and theorists would allow for the feedback necessary to obtain reliable and sufficient data for the use of sophisticated theoretical and statistical methods.

METHODS

In order to benchmark the functionals presented above, we resort to the data set of ref. ⁴³. As a result of our effort to maintain and improve the data set some modifications have been made to the original reference. One entry (BaCu₂GeS₄) was removed from data set due to a discrepancy in the reported experimental structures. Two entries were modified: LiIO₃, which pointed to a structure with same space group but permuted ions, and TiO₂, updated to a more accurate reference. Finally, two new entries have been added, namely BAs and CdO, bringing the total number of entries to 473. Obviously, due to the large size of the data set, these few modifications do not change the conclusions of our previous work.

For each material of the data set, calculations were performed at the corresponding experimental geometry, obtained from the inorganic crystal structure database¹⁵³.

The calculated band gaps were obtained from self-consistent calculations as the difference of the Kohn–Sham eigenvalues at the valence band maximum and conduction band minimum. All calculations were performed within the projector augmented wave formalism¹⁵⁴, as implemented in the Vienna ab initio simulation package (VASP; version 5.4)¹⁵⁵. A custom version of VASP, linked to the library of xc functionals Libxc^{4.5} was used to access xc functionals not implemented in the default distribution. Within the set of pseudopotentials of the VASP distribution, we use the ones recommended by the materials project database¹⁵⁶. All meta-GGA calculations were performed accounting for non-spherical contributions of the density gradient inside the augmentation spheres. We used the same *k*-point sets as in ref. ⁴³, that ensured values converged within 50 meV for both PBE and HLE16 calculations. In addition, we neglect the effect of spin–orbit coupling. Although not entirely negligible, this effect is expected on average to contribute by about 0.1 eV⁴³, which is considerably smaller than the typical average error of approximate functionals.

One final comment is in order. All our calculations were performed with VASP, that uses PAW pseudopotentials to model the electron–ion interaction. The VASP distribution provides a well-tested set of PAW potentials for two functionals (LDA and PBE). This means that most of our calculations, as most of pseudopotential calculations in literature, were performed with functional-inconsistent pseudopotentials. However, the error introduced can be safely neglected if one is only interested in statistical averages¹⁵⁷.

DATA AVAILABILITY

Supplementary material is available on the journal website. All data that support the findings of this study are available in the published article (and in its Supplementary Information files). Source data for figures are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

The library Libxc of exchange–correlation functionals for density-functional theory is freely available online⁵.

Received: 20 December 2019; Accepted: 4 June 2020;

Published online: 10 July 2020

REFERENCES

- Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964).
- Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965).
- Cohen, A. J., Mori-Sánchez, P. & Yang, W. Challenges for density functional theory. *Chem. Rev.* **112**, 289–320 (2012).
- Marques, M. A., Oliveira, M. J. & Burnus, T. Libxc: a library of exchange and correlation functionals for density functional theory. *Comput. Phys. Commun.* **183**, 2272–2281 (2012).
- Lehtola, S., Steigemann, C., Oliveira, M. J. & Marques, M. A. Recent developments in libxc—a comprehensive library of functionals for density functional theory. *SoftwareX* **7**, 1–5 (2018).
- Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- Perdew, J. P., Burke, K. & Ernzerhof, M. Erratum: generalized gradient approximation made simple [Phys. Rev. Lett. 77, 3865 (1996)]. *Phys. Rev. Lett.* **78**, 1396–1396 (1997).
- Becke, A. D. Density-functional thermochemistry. iii. the role of exact exchange. *J. Chem. Phys.* **98**, 5648 (1993).
- 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).
- Heyd, J., Scuseria, G. E. & Ernzerhof, M. Hybrid functionals based on a screened coulomb potential. *J. Chem. Phys.* **118**, 8207 (2003).
- Heyd, J., Scuseria, G. E. & Ernzerhof, M. Erratum: "Hybrid functionals based on a screened Coulomb potential" [J. Chem. Phys. 118, 8207 (2003)]. *J. Chem. Phys.* **124**, 219906 (2006).
- Wiitala, K. W., Hoyer, T. R. & Cramer, C. J. Hybrid density functional methods empirically optimized for the computation of ¹³C and ¹H chemical shifts in chloroform solution. *J. Chem. Theory Comput.* **2**, 1085–1092 (2006).
- Jin, Y. & Bartlett, R. J. Accurate computation of X-ray absorption spectra with ionization potential optimized global hybrid functional. *J. Chem. Phys.* **149**, 064111 (2018).
- Sarmiento-Pérez, R., Botti, S. & Marques, M. A. L. Optimized exchange and correlation semilocal functional for the calculation of energies of formation. *J. Chem. Theory Comput.* **11**, 3844–3850 (2015).
- Tran, F. & Blaha, P. Accurate band gaps of semiconductors and insulators with a semilocal exchange–correlation potential. *Phys. Rev. Lett.* **102**, 226401 (2009).
- Verma, P. & Truhlar, D. G. HLE16: a local Kohn–Sham gradient approximation with good performance for semiconductor band gaps and molecular excitation energies. *J. Phys. Chem. Lett.* **8**, 380–387 (2017).
- Verma, P. & Truhlar, D. G. HLE17: an improved local exchange–correlation functional for computing semiconductor band gaps and molecular excitation energies. *J. Phys. Chem. C* **121**, 7144–7154 (2017).
- Sze, S. M. & Ng, K. K. *Physics of Semiconductor Devices* (Wiley, 2006).
- Neamen, D. A. *Semiconductor Physics and Devices: Basic Principles* (McGraw-Hill, New York, NY, 2012).
- Perdew, J. P. & Levy, M. Physical content of the exact kohn–sham orbital energies: band gaps and derivative discontinuities. *Phys. Rev. Lett.* **51**, 1884–1887 (1983).
- Sham, L. J. & Schlüter, M. Density-functional theory of the energy gap. *Phys. Rev. Lett.* **51**, 1888–1891 (1983).
- Görling, A. Exchange–correlation potentials with proper discontinuities for physically meaningful kohn–sham eigenvalues and band structures. *Phys. Rev. B* **91**, 245120 (2015).
- Godby, R. W., Schlüter, M. & Sham, L. J. Accurate exchange–correlation potential for silicon and its discontinuity on addition of an electron. *Phys. Rev. Lett.* **56**, 2415–2418 (1986).
- Sham, L. J. & Schlüter, M. Density-functional theory of the band gap. *Phys. Rev. B* **32**, 3883–3889 (1985).
- Sharp, R. T. & Horton, G. K. A variational approach to the unipolar many-electron problem. *Phys. Rev.* **90**, 317–317 (1953).
- Talman, J. D. & Shadwick, W. F. Optimized effective atomic central potential. *Phys. Rev. A* **14**, 36–40 (1976).
- Städele, M., Moukara, M., Majewski, J., Vogl, P. & Görling, A. Exact exchange kohn–sham formalism applied to semiconductors. *Phys. Rev. B* **59**, 10031 (1999).
- Neumann, R., Nobes, R. H. & Handy, N. C. Exchange functionals and potentials. *Mol. Phys.* **87**, 1–36 (1996).
- Paier, J. et al. Screened hybrid density functionals applied to solids. *J. Chem. Phys.* **124**, 154709 (2006).
- Giannozzi, P. et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009).
- Sun, J. et al. Self-consistent meta-generalized gradient approximation within the projector-augmented-wave method. *Phys. Rev. B* **84**, 035117 (2011).
- Yao, Y. & Kanai, Y. Plane-wave pseudopotential implementation and performance of scan meta-gga exchange–correlation functional for extended systems. *J. Chem. Phys.* **146**, 224105 (2017).
- Seidl, A., Görling, A., Vogl, P., Majewski, J. A. & Levy, M. Generalized Kohn–Sham schemes and the band-gap problem. *Phys. Rev. B* **53**, 3764–3774 (1996).
- Perdew, J. P. et al. Understanding band gaps of solids in generalized Kohn–Sham theory. *PNAS* **114**, 2801–2806 (2017).
- Cohen, A. J., Mori-Sánchez, P. & Yang, W. Fractional charge perspective on the band gap in density-functional theory. *Phys. Rev. B* **77**, 115123 (2008).
- Mori-Sánchez, P., Cohen, A. J. & Yang, W. Localization and delocalization errors in density functional theory and implications for band-gap prediction. *Phys. Rev. Lett.* **100**, 146401 (2008).
- Mori-Sánchez, P., Cohen, A. J. & Yang, W. Discontinuous nature of the exchange–correlation functional in strongly correlated systems. *Phys. Rev. Lett.* **102**, 066403 (2009).
- Yang, W., Cohen, A. J. & Mori-Sánchez, P. Derivative discontinuity, bandgap and lowest unoccupied molecular orbital in density functional theory. *J. Chem. Phys.* **136**, 204111 (2012).
- Marques, M. A. L., Vidal, J., Oliveira, M. J. T., Reining, L. & Botti, S. Density-based mixing parameter for hybrid functionals. *Phys. Rev. B* **83**, 035119 (2011).
- Salpeter, E. E. & Bethe, H. A. A relativistic equation for bound-state problems. *Phys. Rev.* **84**, 1232–1242 (1951).
- Runge, E. & Gross, E. K. U. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* **52**, 997–1000 (1984).
- Chen, W., Miceli, G., Rignanese, G.-M. & Pasquarello, A. Nonempirical dielectric-dependent hybrid functional with range separation for semiconductors and insulators. *Phys. Rev. Mater.* **2**, 073803 (2018).
- Borlido, P. et al. Large-scale benchmark of exchange–correlation functionals for the determination of electronic band gaps of solids. *J. Chem. Theory Comput.* **15**, 5069–5079 (2019).
- Krukau, A. V., Vydrov, O. A., Izmaylov, A. F. & Scuseria, G. E. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *J. Chem. Phys.* **125**, 224106 (2006).
- Dudarev, S. L., Botton, G. A., Savrasov, S. Y., Humphreys, C. J. & Sutton, A. P. Electron-energy-loss spectra and the structural stability of nickel oxide: an LSDA +U study. *Phys. Rev. B* **57**, 1505–1509 (1998).
- Lichtenstein, A. I., Anisimov, V. I. & Zaanen, J. Density-functional theory and strong interactions: orbital ordering in Mott–Hubbard insulators. *Phys. Rev. B* **52**, R5467–R5470 (1995).
- Doumou, J., Tran, F. & Blaha, P. Limitations of the DFT–1/2 method for covalent semiconductors and transition-metal oxides. *Phys. Rev. B* **99**, 115101 (2019).
- Ferreira, L. G., Marques, M. & Teles, L. K. Approximation to density functional theory for the calculation of band gaps of semiconductors. *Phys. Rev. B* **78**, 125116 (2008).
- Xue, K.-H., Yuan, J.-H., Fonseca, L. R. & Miao, X.-S. Improved LDA-1/2 method for band structure calculations in covalent semiconductors. *Comput. Mater. Sci.* **153**, 493–505 (2018).
- Colonna, N., Nguyen, N. L., Ferretti, A. & Marzari, N. Koopmans–Compliant functionals and potentials and their application to the GW100 test set. *J. Chem. Theory Comput.* **15**, 1905–1914 (2019).
- Das, T., Di Liberto, G., Tosoni, S. & Pacchioni, G. Band gap of 3D metal oxides and quasi-2D materials from hybrid density functional theory: are dielectric-dependent functionals superior? *J. Chem. Theory Comput.* **15**, 6294–6312 (2019).
- Liu, P., Franchini, C., Marsman, M. & Kresse, G. Assessing model-dielectric-dependent hybrid functionals on the antiferromagnetic transition-metal monoxides MnO, FeO, CoO, and NiO. *J. Phys.: Condens. Matter* **32**, 015502 (2020).
- Tsuneda, T. & Hirao, K. Self-interaction corrections in density functional theory. *J. Chem. Phys.* **140**, 18A513 (2014).
- Schmidt, J., Marques, M. R. G., Botti, S. & Marques, M. A. L. Recent advances and applications of machine learning in solid-state materials science. *npj Comput. Mater.* **5**, 1–36 (2019).
- Xie, T. & Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).

56. Isayev, O. et al. Universal fragment descriptors for predicting properties of inorganic crystals. *Nat. Commun.* **8**, 15679 (2017).
57. Rajan, A. C. et al. Machine-learning-assisted accurate band gap predictions of functionalized MXene. *Chem. Mater.* **30**, 4031–4038 (2018).
58. Zhuo, Y., Tehrani, A. M. & Brgoch, J. Predicting the band gaps of inorganic solids by machine learning. *J. Phys. Chem. Lett.* **9**, 1668–1673 (2018).
59. Perdew, J. P. & Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **23**, 5048–5079 (1981).
60. Perdew, J. P. et al. Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008).
61. Becke, A. D. & Johnson, E. R. A simple effective potential for exchange. *J. Chem. Phys.* **124**, 221101 (2006).
62. Sun, J., Ruzsinszky, A. & Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **115**, 036402 (2015).
63. Freysoldt, C. et al. First-principles calculations for point defects in solids. *Rev. Mod. Phys.* **86**, 253–305 (2014).
64. Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: the PBE0 model. *J. Chem. Phys.* **110**, 6158–6170 (1999).
65. Perdew, J. P., Ernzerhof, M. & Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **105**, 9982–9985 (1996).
66. Perdew, J. P. & Schmidt, K. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conf. Proc.* **577**, 1 (2001).
67. Finzel, K. & Baranov, A. I. A simple model for the Slater exchange potential and its performance for solids. *Int. J. Quantum Chem.* **117**, 40–47 (2017).
68. Kohn, W. Density functional and density matrix method scaling linearly with the number of atoms. *Phys. Rev. Lett.* **76**, 3168–3171 (1996).
69. Slater, J. C., Wilson, T. M. & Wood, J. H. Comparison of several exchange potentials for electrons in the Cu⁺ ion. *Phys. Rev.* **179**, 28–38 (1969).
70. Slater, J. C. A Simplification of the Hartree-Fock method. *Phys. Rev.* **81**, 385–390 (1951).
71. Clementi, E. & Roetti, C. Roothaan-Hartree-Fock atomic wavefunctions. *Data Nucl. Data Tables* **14**, 177–478 (1974).
72. Levy, M. & Perdew, J. P. Hellmann-Feynman, virial, and scaling requisites for the exact universal density functionals. Shape of the correlation potential and diamagnetic susceptibility for atoms. *Phys. Rev. A* **32**, 2010–2021 (1985).
73. Engel, E. & Vosko, S. H. Exact exchange-only potentials and the virial relation as microscopic criteria for generalized gradient approximations. *Phys. Rev. B* **47**, 13164–13174 (1993).
74. Ou-Yang, H. & Levy, M. Theorem for exact local exchange potential. *Phys. Rev. Lett.* **65**, 1036–1039 (1990).
75. Tran, F., Blaha, P. & Schwarz, K. Band gap calculations with Becke-Johnson exchange potential. *J. Phys.: Condens. Matter* **19**, 196208 (2007).
76. Tran, F. & Blaha, P. Importance of the kinetic energy density for band gap calculations in solids with density functional theory. *J. Phys. Chem. A* **121**, 3318–3325 (2017).
77. Tran, F. et al. Semilocal exchange-correlation potentials for solid-state calculations: current status and future directions. *J. Appl. Phys.* **126**, 110902 (2019).
78. Armiento, R. & Mattsson, A. E. Functional designed to include surface effects in self-consistent density functional theory. *Phys. Rev. B* **72**, 085108 (2005).
79. Kohn, W. & Mattsson, A. E. Edge electron gas. *Phys. Rev. Lett.* **81**, 3487–3490 (1998).
80. Mattsson, A. E. et al. The AM05 density functional applied to solids. *J. Chem. Phys.* **128**, 084714 (2008).
81. Peverati, R., Zhao, Y. & Truhlar, D. G. Generalized gradient approximation that recovers the second-order density-gradient expansion with optimized across-the-board performance. *J. Phys. Chem. Lett.* **2**, 1991–1997 (2011).
82. Hammer, B., Hansen, L. B. & Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **59**, 7413–7421 (1999).
83. Perdew, J. P. In *Proceedings of the WE-Heraeus-Seminar and 21st Annual International Symposium on Electronic Structure of Solids* (eds Ziesche, P. & Eschrig, H.) 11 (Akademie Verlag, Berlin, 1991).
84. Perdew, J. P. et al. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B* **46**, 6671–6687 (1992).
85. Perdew, J. P. et al. Erratum: Atoms, molecules, solids, and surfaces: applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B* **48**, 4978–4978 (1993).
86. Peverati, R. & Truhlar, D. G. Performance of the M11-L density functional for bandgaps and lattice constants of unary and binary semiconductors. *J. Chem. Phys.* **136**, 134704 (2012).
87. Armiento, R. & Kümmel, S. Orbital localization, charge transfer, and band gaps in semilocal density-functional theory. *Phys. Rev. Lett.* **111**, 036402 (2013).
88. Vlček, V., Steinle-Neumann, G., Leppert, L., Armiento, R. & Kümmel, S. Improved ground-state electronic structure and optical dielectric constants with a semi-local exchange functional. *Phys. Rev. B* **91**, 035107 (2015).
89. Aschbrock, T., Armiento, R. & Kümmel, S. Orbital nodal surfaces: topological challenges for density functionals. *Phys. Rev. B* **95**, 245118 (2017).
90. Aschbrock, T., Armiento, R. & Kümmel, S. Challenges for semilocal density functionals with asymptotically nonvanishing potentials. *Phys. Rev. B* **96**, 075140 (2017).
91. Garhammer, J., Hofmann, F., Armiento, R. & Kümmel, S. On the challenge to improve the density response with unusual gradient approximations. *Eur. Phys. J. B* **91**, 159 (2018).
92. Perdew, J. P. & Wang, Y. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* **45**, 13244 (1992).
93. van Leeuwen, R. & Baerends, E. J. Exchange-correlation potential with correct asymptotic behavior. *Phys. Rev. A* **49**, 2421–2431 (1994).
94. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988).
95. Gaiduk, A. P. & Staroverov, V. N. How to tell when a model Kohn-Sham potential is not a functional derivative. *J. Chem. Phys.* **131**, 044107 (2009).
96. Karolewski, A., Armiento, R. & Kümmel, S. Electronic excitations and the Becke-Johnson potential: the need for and the problem of transforming model potentials to functional derivatives. *Phys. Rev. A* **88**, 052519 (2013).
97. Cerqueira, T. F. T., Oliveira, M. J. T. & Marques, M. A. L. Benchmarking the AK13 exchange functional: ionization potentials and electron affinities. *J. Chem. Theory Comput.* **10**, 5625–5629 (2014).
98. Choi, J., Chang, E., Anstine, D. M., Madjet, M. E.-A. & Chakraborty, H. S. Effects of exchange-correlation potentials on the density-functional description of C 60 versus C 240 photoionization. *Phys. Rev. A* **95**, 023404 (2017).
99. Marques, M. A. L., Castro, A. & Rubio, A. Assessment of exchange-correlation functionals for the calculation of dynamical properties of small clusters in time-dependent density functional theory. *J. Chem. Phys.* **115**, 3006–3014 (2001).
100. Perdew, J. P., Kurth, S., Zupan, A. & Blaha, P. Accurate density functional with correct formal properties: a step beyond the generalized gradient approximation. *Phys. Rev. Lett.* **82**, 2544–2547 (1999).
101. Perdew, J. P., Kurth, S., Zupan, A. & Blaha, P. Erratum: Accurate density functional with correct formal properties: a step beyond the generalized gradient approximation [Phys. Rev. Lett. 82, 2544 (1999)]. *Phys. Rev. Lett.* **82**, 5179–5179 (1999).
102. Adamo, C., Ernzerhof, M. & Scuseria, G. E. The meta-GGA functional: thermochemistry with a kinetic energy density dependent exchange-correlation functional. *J. Chem. Phys.* **112**, 2643–2649 (2000).
103. Tran, F., Stelzl, J. & Blaha, P. Rungs 1 to 4 of DFT Jacob's ladder: extensive test on the lattice constant, bulk modulus, and cohesive energy of solids. *J. Chem. Phys.* **144**, 204120 (2016).
104. Rabuck, A. D. & Scuseria, G. E. Performance of recently developed kinetic energy density functionals for the calculation of hydrogen binding strengths and hydrogen-bonded structures. *Theor. Chem. Acc.* **104**, 439–444 (2000).
105. Tao, J., Perdew, J. P., Staroverov, V. N. & Scuseria, G. E. Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **91**, 146401 (2003).
106. Perdew, J. P., Tao, J., Staroverov, V. N. & Scuseria, G. E. Meta-generalized gradient approximation: explanation of a realistic nonempirical density functional. *J. Chem. Phys.* **120**, 6898–6911 (2004).
107. Perdew, J. P., Ruzsinszky, A., Csonka, G. I., Constantin, L. A. & Sun, J. Workhorse semilocal density functional for condensed matter physics and quantum chemistry. *Phys. Rev. Lett.* **103**, 026403 (2009).
108. Perdew, J. P., Ruzsinszky, A., Csonka, G. I., Constantin, L. A. & Sun, J. Erratum: Workhorse semilocal density functional for condensed matter physics and quantum chemistry [Phys. Rev. Lett. 103, 026403 (2009)]. *Phys. Rev. Lett.* **106**, 179902 (2011).
109. Sun, J., Xiao, B. & Ruzsinszky, A. Communication: effect of the orbital-overlap dependence in the meta generalized gradient approximation. *J. Chem. Phys.* **137**, 051101 (2012).
110. Sun, J., Perdew, J. P. & Ruzsinszky, A. Semilocal density functional obeying a strongly tightened bound for exchange. *Proc. Natl Acad. Sci. USA* **112**, 685–689 (2015).
111. von Weizsäcker, C. F. Zur Theorie der Kernmassen. *Z. Phys.* **96**, 431–458 (1935).
112. Thomas, L. H. The calculation of atomic fields. *Proc. Camb. Philos. Soc.* **23**, 542–548 (1927).
113. Fermi, E. Un metodo statistico per la determinazione di alcune proprietà dell'atomo. *Rend. Accad. Naz. Lincei* **6**, 602–607 (1927).
114. Perdew, J. P., Ruzsinszky, A., Sun, J. & Burke, K. Gedanken densities and exact constraints in density functional theory. *J. Chem. Phys.* **140**, 18A533 (2014).
115. Elliott, P. & Burke, K. Non-empirical derivation of the parameter in the B88 exchange functional. *Can. J. Chem.* **87**, 1485–1491 (2009).

116. Zhang, Y. et al. Efficient first-principles prediction of solid stability: towards chemical accuracy. *npj Comput. Mater.* **4**, 9 (2018).
117. Mezei, P. D., Csonka, G. I. & Kállay, M. Simple modifications of the SCAN meta-generalized gradient approximation functional. *J. Chem. Theory Comput.* **14**, 2469–2479 (2018).
118. Bartók, A. P. & Yates, J. R. Regularized SCAN functional. *J. Chem. Phys.* **150**, 161101 (2019).
119. Furness, J. W. & Sun, J. Enhancing the efficiency of density functionals with an improved iso-orbital indicator. *Phys. Rev. B* **99**, 041119 (2019).
120. Tao, J. & Mo, Y. Accurate semilocal density functional for condensed-matter physics and quantum chemistry. *Phys. Rev. Lett.* **117**, 073001 (2016).
121. Jana, S., Sharma, K. & Samal, P. Improving the performance of tao–mo non-empirical density functional with broader applicability in quantum chemistry and materials science. *J. Phys. Chem. A* **123**, 6356–6369 (2019).
122. Aschebrook, T. & Kümmel, S. Ultranonlocality and accurate band gaps from a meta-generalized gradient approximation. *Phys. Rev. Res.* **1**, 033082 (2019).
123. Zhao, Y. & Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and non-covalent interactions. *J. Chem. Phys.* **125**, 194101 (2006).
124. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120**, 215–241 (2008).
125. Wang, Y., Jin, X., Yu, H. S., Truhlar, D. G. & He, X. Revised M06-L functional for improved accuracy on chemical reaction barrier heights, noncovalent interactions, and solid-state physics. *Proc. Natl Acad. Sci. USA* **114**, 8487–8492 (2017).
126. Peverati, R. & Truhlar, D. G. M11-L: a local density functional that provides improved accuracy for electronic structure calculations in chemistry and physics. *J. Phys. Chem. Lett.* **3**, 117–124 (2012).
127. Peverati, R. & Truhlar, D. G. An improved and broadly accurate local approximation to the exchange–correlation density functional: The MN12-L functional for electronic structure calculations in chemistry and physics. *Phys. Chem. Chem. Phys.* **14**, 13171 (2012).
128. Räsänen, E., Pittalis, S. & Proetto, C. R. Universal correction for the Becke–Johnson exchange potential. *J. Chem. Phys.* **132**, 044112 (2010).
129. Tran, F., Blaha, P. & Schwarz, K. How close are the Slater and Becke–Roussel potentials in solids? *J. Chem. Theory Comput.* **11**, 4717–4726 (2015).
130. Becke, A. D. & Roussel, M. R. Exchange holes in inhomogeneous systems: a coordinate-space model. *Phys. Rev. A* **39**, 3761–3767 (1989).
131. Oliveira, M. J. T., Räsänen, E., Pittalis, S. & Marques, M. A. L. Toward an all-around semilocal potential for electronic exchange. *J. Chem. Theory Comput.* **6**, 3664–3670 (2010).
132. Civalieri, B., Presti, D., Dovesi, R. & Savin, A. in *Chemical Modelling* (ed. Springborg, M.), vol. 9, 168–185 (Royal Society of Chemistry, Cambridge, 2012).
133. Kendall, M. G. A new measure of rank correlation. *Biometrika* **30**, 81–93 (1938).
134. Tran, F., Blaha, P., Betzinger, M. & Blügel, S. Comparison between exact and semilocal exchange potentials: an all-electron study for solids. *Phys. Rev. B* **91**, 165121 (2015).
135. Kuisma, M., Ojanen, J., Enkovaara, J. & Rantala, T. T. Kohn–Sham potential with discontinuity for band gap materials. *Phys. Rev. B* **82**, 115106 (2010).
136. Coskun, D., Jerome, S. V. & Friesner, R. A. Evaluation of the performance of the B3LYP, PBE0, and M06 DFT functionals, and DBLOC-corrected versions, in the calculation of redox potentials and spin splittings for transition metal containing systems. *J. Chem. Theory Comput.* **12**, 1121–1128 (2016).
137. Garza, A. J. & Scuseria, G. E. Predicting band gaps with hybrid density functionals. *J. Phys. Chem. Lett.* **7**, 4165–4170 (2016).
138. Crowley, J. M., Tahir-Kheli, J. & Goddard, W. A. Resolution of the band gap prediction problem for materials design. *J. Phys. Chem. Lett.* **7**, 1198–1203 (2016).
139. Lieb, E. H. & Oxford, S. Improved lower bound on the indirect coulomb energy. *Int. J. Quantum Chem.* **19**, 427 (1981).
140. Zhang, Y. & Yang, W. Comment on “Generalized gradient approximation made simple”. *Phys. Rev. Lett.* **80**, 890–890 (1998).
141. Constantin, L. A., Fabiano, E., Laricchia, S. & Della Sala, F. Semiclassical neutral atom as a reference system in density functional theory. *Phys. Rev. Lett.* **106**, 186406 (2011).
142. Xu, X. & Goddard, W. A. The extended Perdew–Burke–Ernzerhof functional with improved accuracy for thermodynamic and electronic properties of molecular systems. *J. Chem. Phys.* **121**, 4068–4082 (2004).
143. Haas, P. et al. Systematic investigation of a family of gradient-dependent functionals for solids. *Phys. Rev. B* **81**, 125136 (2010).
144. Fabiano, E., Constantin, L. A. & Della Sala, F. Two-dimensional scan of the performance of generalized gradient approximations with Perdew–Burke–Ernzerhof-like enhancement factor. *J. Chem. Theory Comput.* **7**, 3548–3559 (2011).
145. Ouyang, R., Curtarolo, S., Ahmetcik, E., Scheffler, M. & Ghiringhelli, L. M. SISSO: a compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. *Phys. Rev. Mater.* **2**, 083802 (2018).
146. Plumb, G., Molitor, D. & Talwalkar, A. S. Model agnostic supervised local explanations. *Adv. Neural Inform. Proces. Syst.* 2515–2524 (2018).
147. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
148. Lipton, Z. C. The mythos of model interpretability. *Queue* **16**, 30:31–30:57 (2018).
149. Ward, L. et al. Matminer: an open source toolkit for materials data mining. *Comput. Mater. Sci.* **152**, 60–69 (2018).
150. Allen, F. H., Gergert, G. & Sievers, R. (eds.) *Crystallographic Databases* (International Union of Crystallography, Chester, 1987).
151. Zhang, Y. & Ling, C. A strategy to apply machine learning to small datasets in materials science. *npj Comput. Mater.* **4**, 25 (2018).
152. Hu, W., Lin, L., Banerjee, A. S., Vecharynski, E. & Yang, C. Adaptively compressed exchange operator for large-scale hybrid density functional calculations with applications to the adsorption of water on silicene. *J. Chem. Theory Comput.* **13**, 1188–1198 (2017).
153. Kiselyova, N. N., Dudarev, V. A. & Korzhuyev, M. A. Database on the bandgap of inorganic substances and materials. *Inorg. Mater. Appl. Res.* **7**, 34–39 (2016).
154. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
155. Kresse, G. & Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
156. Jain, A. et al. Commentary: The materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
157. Borlido, P., Doumont, J., Tran, F., Marques, M. A. L. & Botti, S. Validation of pseudopotential calculations for the electronic band gap of solids. *J. Chem. Theory Comput.* **15**, 5069–5079 (2020).

ACKNOWLEDGEMENTS

M.A.L.M. and S.B. acknowledge partial support from the DFG through the projects TRR 227, SFB 1375, FOR 2857, BO 4280/8-1, and MA 6787/9-1.

AUTHOR CONTRIBUTIONS

M.A.L.M. and S.B. designed and directed the project. A.W.H. and P.B. performed the benchmark calculations and the reoptimization of the functionals. J.S. developed the machine learning models. All authors contributed to the analysis of the data and the writing and revision of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at <https://doi.org/10.1038/s41524-020-00360-0>.

Correspondence and requests for materials should be addressed to M.A.L.M. or S.B.

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.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020