Performance comparison of r²SCAN and SCAN metaGGA density functionals for solid materials via an automated, high-throughput computational workflow

Ryan Kingsbury [1,2] Ayush S. Gupta [1,3] Christopher J. Bartel [1] Jason M. Munro,[3] Shyam Dwaraknath,[3] Matthew Horton,[3] and Kristin A. Persson[1,4,*]

[1]*Department of Materials Science and Engineering, University of California Berkeley, Berkeley, California 94720, USA*
[2]*Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[3]*Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[4]*Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

Computational materials discovery efforts utilize hundreds or thousands of density functional theory calculations to predict material properties. Historically, such efforts have performed calculations at the generalized gradient approximation (GGA) level of theory due to its efficient compromise between accuracy and computational reliability. However, high-throughput calculations at the higher metaGGA level of theory are becoming feasible. The strongly constrained and appropriately normed (SCAN) metaGGA functional offers superior accuracy to GGA across much of chemical space, making it appealing as a general-purpose metaGGA functional, but it suffers from numerical instabilities that impede its use in high-throughput workflows. The recently developed r²SCAN metaGGA functional promises accuracy similar to SCAN in addition to more robust numerical performance. However, its performance compared to SCAN has yet to be evaluated over a large group of solid materials. In this paper, we compared r²SCAN and SCAN predictions for key properties of approximately 6000 solid materials using a newly developed high-throughput computational workflow. We find that r²SCAN predicts formation energies more accurately than SCAN and PBEsol for both strongly and weakly bound materials and that r²SCAN predicts systematically larger lattice constants than SCAN. We also find that r²SCAN requires modestly fewer computational resources than SCAN and offers significantly more reliable convergence. Thus, our large-scale benchmark confirms that r²SCAN has delivered on its promises of numerical efficiency and accuracy, making it a preferred choice for high-throughput metaGGA calculations.

I. INTRODUCTION

Density functional theory (DFT) has emerged as one of the most widely used computational methods for predicting material properties in recent decades [1,2]. Hundreds of thousands of DFT calculations now populate materials databases, such as the Materials Project [3], Novel Materials Discovery Repository (NOMAD) [4], or the Open Quantum Materials Database (OQMD) [5,6], laying the foundation for a new era of data-driven materials discovery [7].

The vast majority of these calculations employ the Perdew-Burke-Ernzerhof (PBE) [8] generalized gradient approximation (GGA) [9] functional, due to its popularity among researchers and its efficient compromise between high accuracy and high performance across a wide variety of chemistries and properties. However, as a semilocal GGA functional, PBE and its variants, such as PBEsol [10] have well-documented and systematic errors related to electron self-interaction [11] and fail to capture medium- and long-range dispersion [12]. These errors compromise their predictive accuracy for many properties of interest. Specifically, PBE systematically underpredicts the magnitude of formation energies (i.e., underbinds) [13], slightly overpre-

dicts lattice parameters [13,14], and severely underpredicts semiconductor band gaps [13,15,16]. In principle, higher levels of theory such as metaGGA DFT functionals can capture medium-range dispersion interactions and should exhibit smaller self-interaction errors than GGA functionals such as PBE [11]. However, metaGGA functionals have historically been either too specific to selected properties or chemistries and/or too computationally demanding to be feasible for high-throughput calculations across the entire periodic table.

The challenge of generality was addressed by Sun and co-workers via development of the strongly constrained and appropriately normed (SCAN) [17] functional. This nonempirical metaGGA functional has been shown to be substantially more accurate than PBE for predicting lattice constants and ground-state structures of solids [14,18–21] and modestly more accurate for semiconductor band gaps [13,16,22]. SCAN predicts formation energies more accurately than PBE for strongly bound compounds but less accurately than PBE for weakly bound compounds (e.g., intermetallics) [13]. Researchers have also noted that SCAN underpredicts some lattice parameters compared to experiment [13,22] and identified shortcomings in SCAN's predictions of the magnetic moments of ferromagnetic systems [13,22–26], the ground-state polymorphs of selected binary compounds [14], and the bulk moduli of crystalline solids and water ice [22,27].
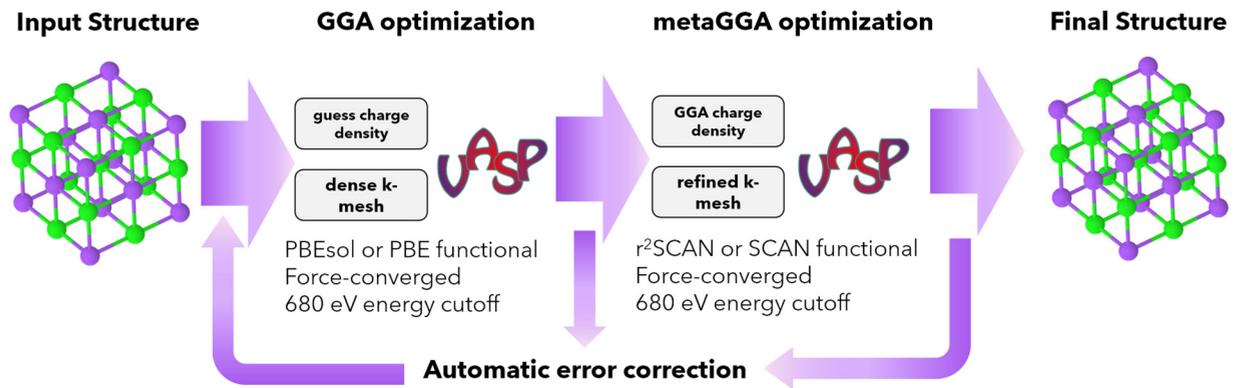
*kapersson@lbl.gov

FIG. 1. Automated workflow for metaGGA calculations. The input structure is relaxed using a GGA functional to construct an initial guess of the charge density. A high k-point density is used in this step. The output charge density from the GGA optimization is used as the initial guess in the subsequent metaGGA optimization using either r^{2}SCAN or SCAN. The band gap estimated from the GGA calculation is used to refine the k-point density for the metaGGA optimization, where metals have the highest k-point density and semiconductors or nonmetals have lower k-point density. Automated error correction routines adjust settings and restart calculations that fail for well-defined reasons, improving reliability.

Nevertheless, the generally superior accuracy of SCAN compared to PBE across many chemical systems and properties makes it appealing as a general-purpose functional for solids. However, it has a higher computational cost than PBE (by a factor of about 5 [13,28]) and suffers from numerical instability. This numerical instability, in particular, makes it impractical to reliably and efficiently perform automated high-throughput calculations with SCAN.

To mitigate the computational challenge, Furness and co-workers recently introduced r^{2}SCAN [29], a modification of the original SCAN functional with substantially improved numerical stability, allowing calculations to converge much more reliably than in the original SCAN functional. This improved stability is achieved by utilizing a smoother switching function to interpolate between the slowly varying and single-orbital density limits, and by relaxing one of the 17 theoretical constraints satisfied by the original SCAN functional (specifically, the fourth-order gradient expansion constraint for exchange GE4X [29,30]). These changes result in a smoother potential-energy surface free of discontinuities. Compared to previous efforts to mitigate numerical difficulties in SCAN [28,31], r^{2}SCAN satisfies a greater number of theoretical constraints and was shown to largely preserve the accuracy of SCAN when benchmarked against several hundred molecular properties and 20 solid lattice constants [29]. Hence, it would appear that the arrival of r^{2}SCAN has, at last, created a general-purpose numerically robust metaGGA functional. However, its accuracy compared to SCAN has yet to be demonstrated for a large group of solid materials.

To further promote progress towards high-throughput metaGGA DFT calculations for solids, in this paper, we compare the formation energies, cell volumes, and electronic structures of approximately 6000 solid materials calculated in r^{2}SCAN and SCAN by means of an automated high-throughput workflow. We show that r^{2}SCAN achieves comparable or even improved accuracy compared to SCAN and reaches convergence much more reliably for this large and diverse set of calculations.

II. METHODOLOGY

A. Automated workflow for metaGGA calculations

DFT calculations were carried out using a two-step workflow comprising an initial GGA structure optimization followed by a structure optimization with the SCAN [17] or r^{2}SCAN [29] metaGGA functionals, coupled with automatic error correction logic (see Fig. 1). The purpose of the initial GGA structure optimization was to generate an initial guess of the structure and charge density at lower computational cost, thereby speeding up the subsequent metaGGA calculation. Performing two optimizations in series also makes the calculation more robust to changes in the size or shape of the unit cell as elaborated further in Sec. S1 of the Supplemental Material [32]. SCAN calculations used PBE [8] for the initial GGA optimization, whereas r^{2}SCAN calculations employed the PBEsol [10] functional, which is a variant of PBE tuned to predict solid lattice constants with greater accuracy. Since the purpose of the initial GGA calculation is simply to accelerate the metaGGA optimization, the final metaGGA result should not be particularly sensitive to the choice of GGA functional.

We employed the Vienna *ab initio* simulation package (VASP) [33,34], version 6.1.1 with custom patches for r^{2}SCAN, in conjunction with projector augmented-wave (PAW) pseudopotentials [35] and a plane-wave energy cutoff of 680 eV. Note that r^{2}SCAN is officially available in VASP as of version 6.2. k-point grids were generated automatically by VASP using KSPACING values ranging from 0.22 to 0.44 Å^{-1}, which were determined from the GGA-estimated band gap of each material based on the work of Wisesa *et al.* [36]. Plane-wave energy cutoff and k-point density settings were selected such that formation energies converged within approximately 1 meV/atom for a benchmark set of 21 materials (listed in Sec. S1 of the Supplemental Material [32]) and were selected to be conservatively high. All calculations used pseudopotentials from the "PBE PAW datasets version 54" set released in September 2015; a list of the specific POTCAR symbols used for each element is provided in Sec. S9 of the Supplemental Material [32]. Although these pseudopotentials were

developed for use with the PBE functional, their use with SCAN is common practice because no SCAN-specific pseudopotentials are available for use in VASP. Additional details related to development of our computational workflow are provided in the Supplemental Material [32] and supporting Refs. [37,38].

B. Selection of materials

The dataset we analyze below includes 6 307 distinct materials, comprising 412 elements, 5 297 binary materials, and 598 ternary materials whose elemental compositions cover the majority of the periodic table (see Fig. S24 of the Supplemental Material [32]). We first screened the Materials Project Database [39] for materials that were within 20 meV of the convex energy hull and had 20 or fewer sites, resulting in a set of approximately 45 000 materials. We retrieved PBE-relaxed structures for each of these from the Materials Project REST API [39], which we used as starting structures in our computational workflow. From this set, we prioritized elements, ground states, and materials close to the convex energy hull and performed metaGGA calculations for as many materials as possible within the computational resources available to our project. We were able to complete approximately 8 000 and 25 000 materials using SCAN and r²SCAN, respectively. We observed considerably more reliable convergence with r²SCAN than with SCAN (as discussed further below), which allowed us to complete calculations for more materials with this functional. Among these SCAN and r²SCAN calculations, there are 6 307 materials (including 5 895 nonelemental solids) for which both SCAN and r²SCAN calculations were completed. Some 5 466 of these materials correspond to structures reported in the Inorganic Crystal Structure Database [40], indicating that they represent experimentally confirmed structures. We use this set of materials to compare the properties predicted by the two functionals. We note that the $\approx 1\,700$ materials for which we completed SCAN but not r²SCAN calculations do not indicate cases where r²SCAN failed to converge. Rather, after completing our SCAN calculations, we chose to prioritize calculations slightly differently for r²SCAN, and, hence, some materials originally completed in SCAN were not attempted in r²SCAN.

III. RESULTS AND DISCUSSION

A. Relative comparison of r²SCAN vs SCAN

1. Formation energy

Computed formation energies predicted by r²SCAN and SCAN are summarized in Fig. 2(a). In an extensive benchmark of the original SCAN functional, Isaacs and Wolverton [13] observed that SCAN formation energies were somewhat more accurate for strongly bound compared to weakly bound materials, where strongly bound materials are those with formation energies $\leqslant -1$ eV/atom and weakly bound materials are those with formation energies between 0 and -1 eV/atom. We adopt the same categories here to facilitate comparison with this prior work. By this definition, our data contain 1 428 and 4 317 strongly and weakly bound materials, respectively. We exclude any materials containing U, Np, or Pu (150 materials) because many exhibited exceptionally large differences in formation energy between the two
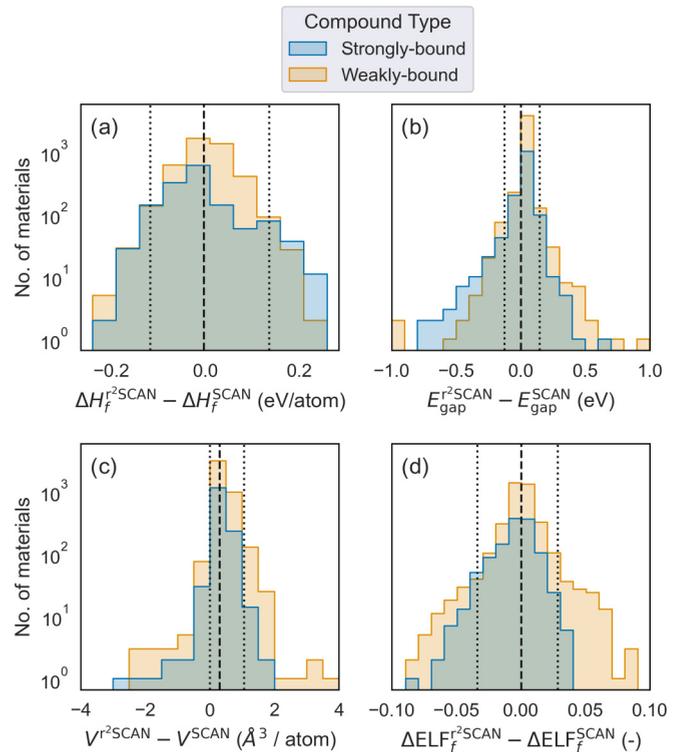

FIG. 2. Changes in (a) formation energy, (b) band gap, (c) cell volume, and (d) formation electron localization function ($\Delta\mathrm{ELF}_f$; see Sec. S5 of the Supplemental Material [32]) when computed in r²SCAN vs SCAN. Note that the y axis is logarithmic. Dashed and dotted vertical lines represent the median differences and two-sided 95th percentile differences, respectively, across both material categories.

functionals. For these cases, we found that r²SCAN-predicted formation energies were substantially more accurate than SCAN-predicted energies compared to experiment. However, since these calculations were performed without spin-orbit coupling, the results must be interpreted with caution, and we do not consider them in detail here. Further discussion is provided in Sec. S4 of the Supplemental Material [32] and supporting Refs. [13,41].

Overall, r²SCAN and SCAN predicted similar formation energies for most materials within both the strongly and weakly bound categories, as indicated by the fact that the median difference in formation energy was only -5 meV/atom. Nevertheless, there were substantial differences for many materials. For 95% of strongly bound materials, $\Delta H_f^{\mathrm{r^2SCAN}}$ differed from $\Delta H_f^{\mathrm{SCAN}}$ by ≈ -135 to $+170$ meV/atom, whereas for weakly bound materials, the 95th-percentile difference in formation energy was ≈ -105 to $+115$ meV/atom. Although the absolute differences in formation energy were similar for strongly and weakly bound materials, in *relative* terms they are much more significant for weakly bound materials since the magnitude of ΔH_f for weakly bound materials is smaller by definition. Reassuringly, we find that despite these apparently large relative differences in predicted ΔH_f for weakly bound materials, r²SCAN has a lower average error compared to experiment by every measure (see Sec. III B 1).

In Sec. S7 of the Supplemental Material [32] we analyze how differences in r^2SCAN and SCAN formation energies relate to specific chemistries. Among strongly and weakly bound materials, the largest positive differences (i.e., materials for which $\Delta H_f^{r^2\mathrm{SCAN}}$ was less negative than $\Delta H_f^{\mathrm{SCAN}}$) were materials containing Pt or Au, while Co-, Ni-, Rh-, and Pd-containing materials also tended to have less negative $\Delta H_f^{r^2\mathrm{SCAN}}$. On the other hand, the largest negative differences in formation energy were observed for Cs-, Pa-, Br-, and Bi-containing materials. It is unclear why the largest differences in formation energy are associated with materials containing these particular elements. The major difference in construction between r^2SCAN and SCAN (the change in switching function and in the formulation of the gradient expansion for exchange) have the greatest effect in regions of slowly varying electron density which would be encountered in metallic compounds [42]. Since the compounds with large positive differences tend to contain (transition) metals, it is possible these elements happen to be particularly sensitive to the different formulation of exchange. For the negative differences, a possible explanation may be that intermediate van der Waals interactions in the elemental Cs, Pa, Br, and Bi phases are less captured by r^2SCAN than SCAN [42]. In general, however, differences in elemental energies do not explain our observations. For pure elements, the difference in electronic energy between r^2SCAN and SCAN grows systematically larger with the atomic number (Fig. S28 of the Supplemental Material [32]), hence, one might expect materials containing heavier elements to exhibit the largest differences in formation energy. This is not what we observed: formation energies of materials show no such systematic trend (Figs. S29 and S30 of the Supplemental Material [32]). Since the formation energy of a material is calculated by subtracting the energies of elemental references from that of the material, it would appear that the sometimes substantial differences in formation energy predicted by r^2SCAN and SCAN are attributable to different energies of the materials rather than the elemental references.

2. Band gaps

GGA DFT functionals are known to systematically and significantly underestimate band gaps, and this shortcoming is only slightly mitigated by SCAN [13]. Nevertheless, it is instructive to examine whether r^2SCAN-predicted band gaps differ substantially from those predicted by SCAN [Fig. 2(b)]. For strongly and weakly bound materials, r^2SCAN band gaps were within ± 0.15 eV of SCAN-predicted band gaps for 95% of materials studied. r^2SCAN was slightly more likely to predict a smaller band gap than SCAN for strongly bound materials and a larger band gap than SCAN for weakly bound materials.

Qualitative agreement in the metallic character of materials predicted by r^2SCAN and SCAN is arguably more relevant than the quantitative band gap predictions. Out of 5 895 materials for which we computed band gaps, there were 73 cases ($\approx 1\%$, listed in Table S4 of the Supplemental Material [32]) in which r^2SCAN predicted metallic character (zero band gap) when SCAN predicted nonmetallic character or vice versa. In six of these cases, the predictions differed

by $\gtrsim 1$ eV. Manual inspection of the density of states (see Sec S2 of the Supplemental Material [32] and supporting Ref. [43]) for the materials with the largest discrepancies indicates that they represent rare cases in which the band occupancies are particularly sensitive to the exchange energy. For example, the material Sb_2F_{13} was a notable outlier where r^2SCAN predicted a metallic material rather than the large-gap insulator predicted by SCAN, due to a small amount of ferromagnetism in the r^2SCAN case shifting the Fermi level into the valence band. Thus, the subtle differences in construction between r^2SCAN and SCAN can occasionally result in large differences in predicted band gaps. However, this example represents a fictitious unphysical material since the originating crystal structure file was found to have omitted hydrogens when compared against the original publication. We emphasize that we observed these large discrepancies in only $\approx 0.1\%$ of materials in our dataset and that they may be partially attributable to recently identified changes in the way VASP computes the Fermi level.

3. Lattice volumes

r^2SCAN systematically predicted larger lattice volumes for many materials than SCAN, and this systematic difference was observed to a similar extent across both material categories [Fig. 2(c)]. Specifically, the median volumes per atom predicted by r^2SCAN were 1.4% and 1.8%, (0.2–0.3 Å^3/atom) larger than the SCAN predicted volumes for strongly bound and weakly bound materials, respectively. The systematically larger lattice volumes predicted by r^2SCAN compared to SCAN may be fortuitous since SCAN was previously shown to underpredict experimental lattice volumes by an average of 0.11 Å^3/atom [13]. We will examine the accuracy of SCAN and r^2SCAN lattice volumes compared to experiment in a later section.

4. Electron localization

To evaluate the consistency between r^2SCAN and SCAN in a more general way, we next present differences in the electronic structure predicted by the two functionals. Both r^2SCAN and SCAN incorporate information about the kinetic-energy density into their calculation of the exchange and correlation energies by means of the iso-orbital indicator $\alpha = \frac{\tau - \tau_\mathrm{W}}{\tau_\mathrm{unif}}$ or $\bar{\alpha} = \frac{\tau - \tau_\mathrm{W}}{\tau_\mathrm{unif} + \eta \tau_\mathrm{W}}$ for SCAN and r^2SCAN, respectively, where τ is the positive kinetic-energy density, τ_W and τ_unif are the limiting kinetic energies of a single orbital and uniform electron gas, respectively, and $\eta = 0.001$ is a regularization parameter [11,17,29]. The calculated value of the exchange and correlation energies depends on the value of the iso-orbital indicator and, hence, on the bonding regime (e.g., localized/covalent or delocalized/metallic). This ability to adjust for different local electronic environments is a major reason for the superior accuracy of r^2SCAN and SCAN compared to GGA [13] and explains why SCAN requires much smaller Hubbard U values than GGA functionals to accurately predict formation energies of transition-metal oxides [44].

r^2SCAN differs from SCAN primarily in that: (1) it uses $\bar{\alpha}$ instead of α as the iso-orbital indicator (see the definitions above), and (2) it uses a different switching function to adjust the value of the exchange-correlation enhancement factor F_xc,

for different values of $\bar{\alpha}$ [29]. The iso-orbital indicator is closely related to the electron localization function (ELF), which is equal to $(1 + \alpha^2)^{-1}$ [17,29] and ranges from 0 to 1 where a value of 0.5 corresponds to an electron gas, and 1 corresponds to highly localized (covalent) bonding [45,46]. Since both α and $\bar{\alpha}$ and, hence, the value of F_{xc} are directly related to ELF, in Fig. 2(d), we present the difference in the "formation ELF," $\Delta\mathrm{ELF}_f$, predicted by r²SCAN and SCAN. $\Delta\mathrm{ELF}_f$ is calculated by analogy to the formation energy (see Sec. S5 of the Supplemental Material [32]) and represents the degree to which the average amount of electron localization around each atom in a material differs from that in the corresponding elemental references, as determined by Bader analysis [47,48]. By construction, the formation ELF must fall between 0 and 1. However, because high electron localization occurs only in a relatively small fraction of the volume occupied by a crystal (e.g., near the nuclei or along a covalent bond), average values for an entire atomic basin are typically small. An example illustrating the local value of ELF in a crystal structure is provided in Sec S5 in the Supplemental Material [32]. In our dataset, the median values of $\Delta\mathrm{ELF}_f$ were 0.034 and 0.036 (dimensionless) for r²SCAN and SCAN, respectively. These values are 10–20% larger than the median $\Delta\mathrm{ELF}_f$ calculated by PBEsol (0.030), indicating that the two metaGGA functionals predict larger changes in electronic structure during compound formation than GGA.

Figure 2(d) shows that broadly speaking, $\Delta\mathrm{ELF}_f$ values predicted by r²SCAN and SCAN are similar for both categories of materials (as indicated by the differences being centered around 0). There is a slight skew towards r²SCAN predicting smaller $\Delta\mathrm{ELF}_f$ (i.e., less change in localization between elements and compounds) than SCAN with differences of -0.033 to $+0.028$ defining the 95th percentile of all materials. However, in relative terms (i.e., $|\frac{\Delta\mathrm{ELF}_f^{\mathrm{r^2SCAN}} - \Delta\mathrm{ELF}_f^{\mathrm{SCAN}}}{\Delta\mathrm{ELF}_f^{\mathrm{SCAN}}}|$) these changes were quite large with a median relative change of 16% and a 95th percentile relative change of 82%. Hence, whereas r²SCAN and SCAN predict similar $\Delta\mathrm{ELF}_f$ in aggregate, subtle absolute differences in their respective prediction of ELF may result in large relative changes for specific materials.

By definition, the difference in DFT energy between the two functionals when evaluated on the same density (and Kohn-Sham orbitals) is equal to the difference in their exchange and correlation energies, which is a complex function of not only the ELF, but also the density, density gradient, and kinetic-energy density [11,29] which are themselves products of previous self-consistent iterations. Hence, it is difficult to relate changes in $\Delta\mathrm{ELF}_f$ directly to changes in DFT or formation energy. Figure 2(a) suggests that the large relative shifts in $\Delta\mathrm{ELF}_f$ that we observe between r²SCAN and SCAN do not lead to commensurately large shifts in ΔH_f. For comparison, the median and the 95th percentile relative changes in the electronic energy (i.e., the DFT energy) from SCAN to r²SCAN were 26% and 32%, respectively, whereas the median and the 95th percentile changes in ΔH_f were 6.5% and 54%. Nevertheless, Figure 2(d) shows that differences in $\Delta\mathrm{ELF}_f$ tended to be larger for weakly bound materials compared to strongly bound materials, and this fact could be

related to the larger relative changes in formation energy that we observed for weakly bound materials (see above).

B. Experimental benchmarks

1. Formation energy

Having examined how material properties predicted by r²SCAN differ from those of SCAN, we now turn our attention to how accurately r²SCAN and SCAN predict experimental formation energies, volumes, and band gaps. Predictions by the PBEsol GGA functional [10], executed with the same settings as the metaGGA step of the automated workflow (see Sec. II A) are shown as an additional point of comparison. We note that since experimental benchmarking was not the primary objective of this paper, the materials we evaluate in this section are dictated primarily by the calculations we generated rather than through deliberate selection. Complementary efforts are underway by other research groups [42] to benchmark r²SCAN against experimental data using carefully curated sets of materials presented in previous studies [13,18].

Figure 3(a) shows the mean absolute error (MAE) in formation energy for 986 materials. Experimental energies for these materials were obtained from the `expt_formation_enthalpy_kingsbury` dataset distributed with `Matminer` [49], which associates formation energies compiled from Refs. [50–54] with specific crystal structures from the ICSD [40], allowing us to match our computed data with high confidence. Additional details about the dataset are provided in Sec. S8 of the Supplemental Material [32].

For the majority of materials studied, the MAE in formation energy predicted by metaGGA functionals was $\approx$ 80–120 meV/atom. Strikingly, r²SCAN formation energies had approximately 20% and 15% lower MAEs than SCAN for strongly and weakly bound materials, respectively, even though r²SCAN is less theoretically exact (i.e., r²SCAN relaxes the fourth-order gradient expansion constraint for exchange that is satisfied by SCAN [29]). Although surprising from a theoretical standpoint, other recent studies have also reported greater accuracy of r²SCAN compared to SCAN. For example, r²SCAN was found to predict cohesive energies and bulk moduli of solids more accurately than SCAN [55].

We observed in Sec. III A that despite many materials having similar formation energies, there were a number of outlying materials for which $|\Delta H_f^{\mathrm{r^2SCAN}} - \Delta H_f^{\mathrm{SCAN}}| \gtrsim$ 100–200 meV/atom or more. To evaluate the implications of such large differences, we examined the accuracy vs experiment of all strongly or weakly bound materials for which $|\Delta H_f^{\mathrm{r^2SCAN}} - \Delta H_f^{\mathrm{SCAN}}| \geqslant 50$ meV/atom. There are 665 such materials in the entire dataset of which we have experimental data for 345. Among this group of outliers, $\Delta H_f^{\mathrm{r^2SCAN}}$ is more accurate than $\Delta H_f^{\mathrm{SCAN}}$. Specifically, the mean error for r²SCAN was -43.8 meV/atom, wheres the mean error for SCAN was -55.8 meV/atom, and the MAEs were 90.2 and 134.9 meV/atom for r²SCAN and SCAN, respectively.

In addition, we find that for strongly bound materials, both r²SCAN and SCAN predicted formation energy much more accurately than PBEsol. This finding is similar to the
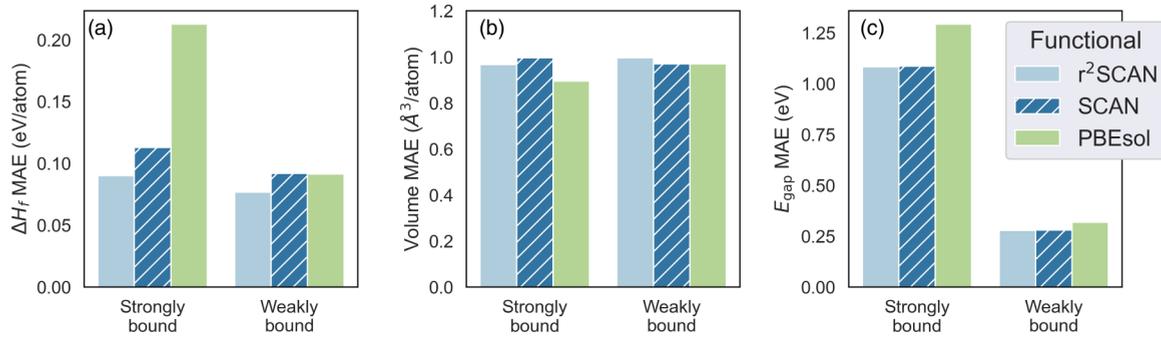
FIG. 3. Mean absolute error compared to experiment in (a) formation energy ($n = 986$ materials), (b) cell volume ($n = 4\,974$ materials), and (c) band-gap ($n = 582$ materials) computed with r^{2}SCAN, SCAN, or PBEsol.

previous observation by by Isaacs and Wolverton [13] that SCAN outperforms PBE for strongly bound materials. For weakly bound materials, however, SCAN was slightly less accurate than PBEsol (also consistent with previous findings [13]), whereas r^{2}SCAN was more accurate. Hence, r^{2}SCAN predicted formation energy more accurately than SCAN or PBEsol for all material categories. This is a fortuitous result: in addition to enhancing numerical stability, it appears that the greater smoothness of the r^{2}SCAN potential energy surface improves accuracy in one of the few areas in which SCAN was less accurate than PBE, as elaborated by Furness *et al.* [30] and Kaplan *et al.* [42].

2. Lattice volume

Figure 3(b) summarizes the performance of the three functionals for predicting cell volume using experimental data obtained from the Inorganic Crystal Structure Database [40]. For strongly bound materials, PBEsol has the lowest MAE of 0.89 Å^3/atom, followed by r^{2}SCAN (0.97 Å^3/atom) and SCAN (1.0 Å^3/atom). For weakly bound materials, PBEsol and SCAN predict volume with a similar MAE of 0.97 Å^3/atom, whereas r^{2}SCAN has a slightly higher MAE of 1.0 Å^3/atom. Overall, neither metaGGA functional shows a clear and significant improvement in lattice volume prediction compared to PBEsol. Although surprising considering that SCAN lattice constants were shown to be more accurate than PBE (albeit underpredicted, whereas PBE lattice constants were overpredicted) [13], it is important to remember that PBEsol was developed specifically to reproduce solid lattice constants with high accuracy.

3. Band gap

The errors in predicted band gap are shown in Fig. 3(c) for 582 materials that were present in both our dataset and the experimental `expt_gap_kingsbury` dataset in `Matminer` [49]. Compilation of the band gap data is described in more detail in Sec. S8 of the Supplemental Material [32].

For strongly bound materials, the MAE in the predicted band gap was nearly identical between r^{2}SCAN and SCAN at 1.078 and 1.081 eV, respectively. The same was true for weakly bound materials; although, in this case, the MAE was much lower at ≈ 0.28 eV. Although the MAEs for both metaGGA functionals were considerable, they were ≈ 0.21 and 0.04 eV lower than the PBEsol MAE for strongly and weakly bound materials, respectively.

C. Computational performance and reliability

Finally, we used the large amount of computed data we generated to develop a qualitative understanding of the relative computational demands and reliability of r^{2}SCAN and SCAN. In Figure 4(a) we present the relative performance of r^{2}SCAN, SCAN, and PBEsol in terms of: (1) total CPU time, (2) total number of ionic steps, and (3) total number of self-consistent field (SCF) cycles (summed over all ionic steps) required to reach convergence. We note that this was not a rigorous computational benchmark because the starting structures for the GGA and metaGGA stages of the workflow were not identical. As described in Sec. II A, each starting structure was optimized using equivalent VASP settings with both: (1) PBE followed by SCAN and (2) PBEsol followed by r^{2}SCAN. All calculations were carried out on the Cori supercomputer at the National Energy Research Scientific Computing Center (Berkeley, CA); however, the parallelization settings (i.e., number of nodes, cores, and multiprocessing tasks) varied slightly among calculations. As such, the information in Fig. 4 should not be considered a definitive representation of the relative computational demands of these functionals but rather a qualitative representation of their performance over a large and diverse set of materials.

As shown in the figure, r^{2}SCAN required approximately 0.5-2$\times$ the CPU time to converge as SCAN, and 0.5-1$\times$ the number of ionic and electronic steps. Compared to PBEsol, r^{2}SCAN required 2-4$\times$ the CPU time, 1-3$\times$ the number of ionic steps, and 1-2$\times$ the number of total SCF cycles. Our results here are largely consistent with a previous rigorous benchmark for CPU time based on single-point calculations, which showed that r^{2}SCAN required approximately 0.9$\times$ the CPU time as SCAN and approximately 4$\times$ as much as PBE [56]. Our finding that fewer ionic steps are required is consistent with a study by Ning and co-workers [57] which found that the smoother potential energy surface of r^{2}SCAN facilitated convergence compared to SCAN. Thus, modestly fewer ionic steps and modestly less CPU time are required to converge r^{2}SCAN than SCAN, but both functionals require considerably more computational resources than PBEsol.

Time required for convergence does not tell the whole story, however, because the data in Fig. 4(a) reflect only *successful* calculations, and, hence, mask the much more reliable convergence of r^{2}SCAN. Among approximately 11 000 and 25 000 total calculations we attempted with SCAN and
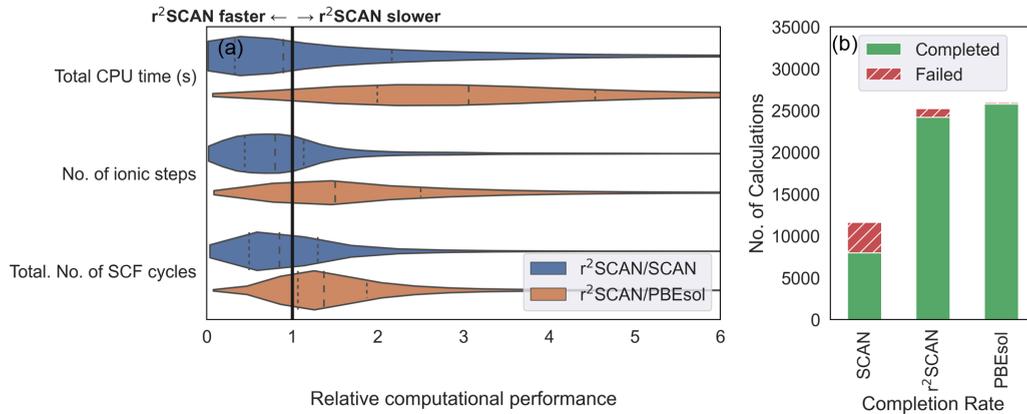
FIG. 4. (a) Relative computational performance of r^{2}SCAN compared to SCAN (blue) or PBEsol (orange), demonstrating that computational time required for r^{2}SCAN is smaller than SCAN but larger than PBEsol. Each variable is plotted as the ratio of the value in the r^{2}SCAN calculation divided by the value in the corresponding SCAN or PBEsol calculation. Dashed lines inside each violin represent the quartiles of the distribution. (b) Completion rate of calculations carried out using each density functional.

r^{2}SCAN, respectively, we observed a completion rate of 96% for r^{2}SCAN but only 69% for SCAN [Fig. 4(b)]. Among the incomplete SCAN calculations, we estimate that at least 25% failed due to unrecoverable errors, whereas the remaining calculations may have simply run out of wall time, which was limited to 48 h in our work. Again, although these failure statistics do not represent a rigorous comparison of the two functionals (for example, the SCAN convergence rate may improve somewhat with very long wall time), they qualitatively highlight the generally much more reliable convergence of r^{2}SCAN compared to SCAN, which was a key objective of its development [29]. Furthermore, it is noteworthy that for weakly bound compounds, PBEsol consumed less than half the computational resources of r^{2}SCAN, while predicting only modestly less accurate formation energies. Hence, PBEsol may still be considered an excellent choice for computing many properties of interest.

IV. SUMMARY AND OUTLOOK

To summarize, we have compared r^{2}SCAN and SCAN predictions for key properties of approximately 6 000 solid materials. We find that r^{2}SCAN predicts substantially similar formation energies, band gaps, and degrees of electron localization as the original SCAN functional, but predicts systematically larger lattice constants. r^{2}SCAN is found to predict formation energies more accurately than SCAN and PBEsol for both strongly and weakly bound materials, while r^{2}SCAN and SCAN calculated band gaps are virtually identical and modestly more accurate than those predicted by PBEsol. For materials containing U, Np, or Pu, r^{2}SCAN predicts formation energies that are substantially different from and considerably more accurate than those predicted by SCAN. The reason for this is not clear and could arise from a fortuitous cancellation of errors related to the lack of spin-

orbit coupling in our calculations and/or as a consequence of the smoother potential-energy surface generated by r^{2}SCAN. With respect to computational reliability, we find that r^{2}SCAN requires modestly fewer computational resources than SCAN but offers much more reliable convergence. Thus, our large-scale benchmark confirms that r^{2}SCAN has delivered on its promises of numerical efficiency and accuracy [29], making it an ideal choice for high-throughput metaGGA calculations.

V. DATA AVAILABILITY

All data referenced herein are publicly available on FIGSHARE [58] and will be integrated into the Materials Project database [39] in the near future. Our computational workflow has been implemented into the PYMATGEN [59], CUSTODIAN [59], and ATOMATE [60] packages as of versions 2020.12.3, 2021.1.8, and 0.9.6, respectively, for readers wishing to utilize it in their own work.

ACKNOWLEDGMENTS

[1] J. Hafner, C. Wolverton, and G. Ceder, Toward computational materials design: The impact of density functional theory on materials research, MRS Bulletin **31**, 659 (2006).

[2] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, The high-throughput highway to computational materials design, Nat. Mater. **12**, 191 (2013).

[3] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, The Materials Project: A materials genome approach to accelerating materials innovation, APL Mater. **1**, 011002 (2013).

[4] C. Draxl and M. Scheffler, Nomad: The FAIR concept for big data-driven materials science, MRS Bull. **43**, 676 (2018).

[5] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, Materials design and discovery with high-throughput density functional theory: The Open Quantum Materials Database (OQMD), JOM **65**, 1501 (2013).

[6] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, and C. Wolverton, The Open Quantum Materials Database (OQMD): Assessing the accuracy of DFT formation energies, npj Comput. Mater. **1**, 15010 (2015).

[7] L. Himanen, A. Geurts, A. S. Foster, and P. Rinke, Data-driven materials science: Status, challenges, and perspectives, Adv. Sci. **6**, 1900808 (2019).

[8] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, Phys. Rev. Lett. **77**, 3865 (1996).

[9] D. C. Langreth and J. P. Perdew, Theory of nonuniform electronic systems. I. analysis of the gradient approximation and a generalization that works, Phys. Rev. B **21**, 5469 (1980).

[10] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces, Phys. Rev. Lett. **100**, 136406 (2008).

[11] J. P. Perdew, Climbing the ladder of density functional approximations, MRS Bull. **38**, 743 (2013).

[12] S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth, Dispersion-Corrected Mean-Field Electronic Structure Methods, Chem. Rev. **116**, 5105 (2016).

[13] E. B. Isaacs and C. Wolverton, Performance of the strongly constrained and appropriately normed density functional for solid-state materials, Phys. Rev. Materials **2**, 063801 (2018).

[14] J. H. Yang, D. A. Kitchaev, and G. Ceder, Rationalizing accurate structure prediction in the meta-GGA SCAN functional, Phys. Rev. B **100**, 035132 (2019).

[15] S. Lany and A. Zunger, Assessment of correction methods for the band-gap problem and for finite-size effects in supercell defect calculations: Case studies for ZnO and GaAs, Phys. Rev. B **78**, 235104 (2008).

[16] Y. Si, M. Li, Z. Zhou, M. Liu, and O. Prezhdo, Improved description of hematite surfaces by the SCAN functional, J. Chem. Phys. **152**, 024706 (2020).

[17] J. Sun, A. Ruzsinszky, and J. P. Perdew, Strongly Constrained and Appropriately Normed Semilocal Density Functional, Phys. Rev. Lett. **115**, 036402 (2015).

[18] Y. Zhang, D. A. Kitchaev, J. Yang, T. Chen, S. T. Dacek, R. A. Sarmiento-Pérez, M. A. Marques, H. Peng, G. Ceder, J. P. Perdew, and J. Sun, Efficient first-principles prediction of solid stability: Towards chemical accuracy, npj Comput. Mater. **4**, 9 (2018).

[19] W. Sun, A. Holder, B. Orvañanos, E. Arca, A. Zakutayev, S. Lany, and G. Ceder, Thermodynamic Routes to Novel Metastable Nitrogen-Rich Nitrides, Chem. Mater. **29**, 6936 (2017).

[20] J.-S. Park, Comparison study of exchange-correlation functionals on prediction of ground states and structural properties, Current Appl. Phys. **22**, 61 (2021).

[21] C. J. Bartel, A. W. Weimer, S. Lany, C. B. Musgrave, and A. M. Holder, The role of decomposition reactions in assessing first-principles predictions of solid stability, npj Comput. Mater. **5**, 4 (2019).

[22] S. Jana, A. Patra, and P. Samal, Assessing the performance of the Tao-Mo semilocal density functional in the projector-augmented-wave method, J. Chem. Phys. **149**, 044120 (2018).

[23] Y. Fu and D. J. Singh, Applicability of the Strongly Constrained and Appropriately Normed Density Functional to Transition-Metal Magnetism, Phys. Rev. Lett. **121**, 207201 (2018).

[24] M. Ekholm, D. Gambino, H. J. M. Jönsson, F. Tasnádi, B. Alling, and I. A. Abrikosov, Assessing the scan functional for itinerant electron ferromagnets, Phys. Rev. B **98**, 094413 (2018).

[25] A. H. Romero and M. J. Verstraete, From one to three, exploring the rungs of jacob's ladder in magnetic alloys, Eur. Phys. J. B **91**, 193 (2018).

[26] D. Mejía-Rodríguez and S. B. Trickey, Analysis of over-magnetization of elemental transition metal solids from the SCAN density functional, Phys. Rev. B **100**, 041113(R) (2019).

[27] J. S. Rego and M. de Konig, Density-functional theory prediction of the elastic constants of ice I h, J. Chem. Phys. **152**, 084502 (2020).

[28] D. Mejia-Rodriguez and S. B. Trickey, Deorbitalized meta-GGA exchange-correlation functionals in solids, Phys. Rev. B **98**, 115161 (2018).

[29] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, Accurate and numerically efficient r2SCAN meta-generalized gradient approximation, J. Phys. Chem. Lett. **11**, 8208 (2020).

[30] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, Construction of meta-GGA functionals through restoration of exact constraint adherence to regularized scan functionals, arXiv:2110.00647.

[31] A. P. Bartók and J. R. Yates, Regularized SCAN functional, J. Chem. Phys. **150**, 161101 (2019).

[32] See Supplemental Material at https://link.aps.org/supplemental/10.1103/PhysRevMaterials.6.013801 for further details concerning workflow development, cases of qualitative disagreement in metallic character, elemental representation among studied materials, anomalous behavior of materials containing U, Np, or Pu, analysis of electron localization, differences in electronic and formation energy of r2SCAN vs SCAN by element, experimental benchmarking data, and list of VASP pseudopotentials used.

[33] G. Kresse and J. Furthmüller, Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set, Comput. Mater. Sci. **6**, 15 (1996).

[34] G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).

[35] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).

[36] P. Wisesa, K. A. McGill, and T. Mueller, Efficient generation of generalized Monkhorst-Pack grids through the use of informatics, Phys. Rev. B **93**, 155109 (2016).

[37] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, and O. Levy, Aflowlib.org: A distributed materials properties repository from high-throughput *ab initio* calculations, Comput. Mater. Sci. **58**, 227 (2012).

[38] K. Choudhary and F. Tavazza, Convergence and machine learning predictions of Monkhorst-Pack k-points and plane-wave cut-off in high-throughput DFT calculations, Comput. Mater. Sci. **161**, 300 (2019).

[39] S. P. Ong, S. Cholia, A. Jain, M. Brafman, D. Gunter, G. Ceder, and K. A. Persson, The materials application programming interface (API): A simple, flexible and efficient API for materials data based on REpresentational state transfer (REST) principles, Comput. Mater. Sci. **97**, 209 (2015).

[40] I. Levin, NIST Inorganic Crystal Structure Database (ICSD), NIST Standard Reference Database Number 3, National Institute of Standards and Technology, Gaithersburg MD, 20899 (2020), https://doi.org/10.18434/M32147.

[41] S. Lany, Semiconductor thermochemistry in density functional calculations, Phys. Rev. B **78**, 245207 (2008).

[42] A. D. Kaplan, J. W. Furness, L. Hou, J. Sun, and J. P. Perdew (private communication).

[43] A. M. Ganose, A. J. Jackson, and D. O. Scanlon, sumo: Command-line tools for plotting and analysis of periodic *ab initio* calculations, J. Open Source Software **3**, 717 (2018).

[44] O. Y. Long, G. Sai Gautam, and E. A. Carter, Evaluating optimal *u* for 3*d* transition-metal oxides within the scan+*u* framework, Phys. Rev. Mater. **4**, 045401 (2020).

[45] P. Fuentealba, E. Chamorro, and J. C. Santos, Understanding and using the electron localization function, *Theoretical and Computational Chemistry* (Elsevier, 2007), Vol. 19, Chap. 5, pp. 57–85.

[46] B. Silvi and A. Savin, Classification of chemical bonds based on topological analysis of electron localization functions, Nature (London) **371**, 683 (1994).

[47] M. Yu and D. R. Trinkle, Accurate and efficient algorithm for bader charge integration, J. Chem. Phys. **134**, 064111 (2011).

[48] W. Tang, E. Sanville, and G. Henkelman, A grid-based bader analysis algorithm without lattice bias, J. Phys.: Condens. Matter **21**, 084204 (2009).

[49] L. Ward, A. Dunn, A. Faghaninia, N. E. Zimmermann, S. Bajaj, Q. Wang, J. Montoya, J. Chen, K. Bystrom, M. Dylla, K. Chard, M. Asta, K. A. Persson, G. J. Snyder, I. Foster, and A. Jain, Matminer: An open source toolkit for materials data mining, Comput. Mater. Sci. **152**, 60 (2018).

[50] M. Chase, *NIST-JANAF Thermochemical Tables*, 4th ed. (American Institute of Physics, 1998).

[51] O. Kubaschewski, C. Alcock, and P. Spencer, *Materials Thermochemistry*, 6th ed. (Pergamon, Oxford, 1993).

[52] G. Kim, S. V. Meschel, P. Nash, and W. Chen, Experimental formation enthalpies for intermetallic phases and other inorganic compounds, Sci. Data **4**, 170162 (2017).

[53] G. Kim, S. Meschel, P. Nash, and W. Chen, "Experimental formation enthalpies for intermetallic phases and other inorganic compounds", figshare repository, https://doi.org/10.6084/m9.figshare.c.3822835.v1 (2017).

[54] A. Wang, R. Kingsbury, M. McDermott, M. Horton, A. Jain, S. P. Ong, S. Dwaraknath, and K. Persson, A framework for quantifying uncertainty in DFT energy corrections, Sci. Rep. **11**, 15496 (2021).

[55] D. Mejía-Rodríguez and S. B. Trickey, Meta-GGA performance in solids at almost GGA cost, Phys. Rev. B **102**, 121109(R) (2020).

[56] D. Mejía-Rodríguez and S. B. Trickey, Spin-Crossover from a Well-Behaved, Low-Cost meta-GGA Density Functional, J. Phys. Chem. A **124**, 9889 (2020).

[57] J. Ning, J. W. Furness, and J. Sun, Reliable lattice dynamics from an efficient density functional, arXiv:2107.11850.

[58] R. Kingsbury, A. Gupta, C. J. Bartel, J. M. Munro, S. Dwaraknath, M. Horton, and K. Persson, "r2SCAN, SCAN, and PBEsol calcluations for 6,307 solid materials", figshare repository, https://doi.org/10.6084/m9.figshare.16564815.v2 (2021).

[59] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis, Comput. Mater. Sci. **68**, 314 (2013).

[60] K. Mathew, J. H. Montoya, A. Faghaninia, S. Dwarakanath, M. Aykol, H. Tang, I. Chu, T. Smidt, B. Bocklund, M. Horton, J. Dagdelen, B. Wood, Z.-K. Liu, J. Neaton, S. P. Ong, K. Persson, and A. Jain, Atomate: A high-level interface to generate, execute, and analyze computational materials science workflows, Comput. Mater. Sci. **139**, 140 (2017).