npj computational materials

www.nature.com/npjcompumats

ARTICLE OPEN

Check for updates

A flexible and scalable scheme for mixing computed formation energies from different levels of theory

Ryan S. Kingsbury[1,2], Andrew S. Rosen[1,3,4], Ayush S. Gupta[1], Jason M. Munro[3], Shyue Ping Ong[3,5], Anubhav Jain[2], Shyam Dwaraknath[3], Matthew K. Horton[3] and Kristin A. Persson[1,6]✉

Computational materials discovery efforts are enabled by large databases of properties derived from high-throughput density functional theory (DFT), which now contain millions of calculations at the generalized gradient approximation (GGA) level of theory. It is now feasible to carry out high-throughput calculations using more accurate methods, such as meta-GGA DFT; however recomputing an entire database with a higher-fidelity method would not effectively leverage the enormous investment of computational resources embodied in existing (GGA) calculations. Instead, we propose here a general procedure by which higher-fidelity, low-coverage calculations (e.g., meta-GGA calculations for selected chemical systems) can be combined with lower-fidelity, high-coverage calculations (e.g., an existing database of GGA calculations) in a robust and scalable manner. We then use legacy PBE(+*U*) GGA calculations and new r^{2}SCAN meta-GGA calculations from the Materials Project database to demonstrate that our scheme improves solid and aqueous phase stability predictions, and discuss practical considerations for its implementation.

npj Computational Materials (2022)8:195 ; https://doi.org/10.1038/s41524-022-00881-w

INTRODUCTION

The advent of large databases of computed material properties, such as the Materials Project[1], AFLOW[2], the Open Quantum Materials Database (OQMD)[3,4], and the Joint Automated Repository for Various Integrated Simulations (JARVIS)[5], has paved the way for a new era of data-driven materials science[6]. These databases now contain computed properties derived from millions of individual calculations, the vast majority of which employ density functional theory (DFT) due to its efficient compromise between computational cost and accuracy. For example, the Materials Project[1] contains computed formation energies for more than 140,000 materials calculated using the Perdew–Burke–Ernzerhof (PBE)[7] generalized gradient approximation (GGA)[8] functional, with a Hubbard *U*, value[9,10] and empirical energy corrections[11] applied to some chemical systems. This data is widely used in machine learning and computational materials screening efforts in which the thermodynamic (meta)stablility of a material is often among the first selection criteria[12].

Despite its versatility and historical success, PBE has well-documented systematic errors related to electron self-interaction[13,14] that are particularly notable in diatomic gases[15,16] and transition metal compounds with localized electronic states[1,16,17]. PBE also fails to capture medium- and long-range dispersion interactions[18], which are important for describing the properties of weakly-bound systems. Even when adjusted using empirical correction schemes, the mean absolute error (MAE) in formation energies predicted by this level of theory is still on the order of 50–200 meV per atom[11,15,16,19–23], although the error in energy differences among polymorphs is typically lower (e.g., 25 meV per atom)[24].

Today, more than a decade after most materials databases were established[6], theoretical advances and growth in computing power have made it feasible to compute large numbers of formation energies at higher levels of theory[19], which could substantially increase their accuracy. For example, we recently showed that the restored-regularized Strongly Constrained and Appropriately Normed (r^{2}SCAN) meta-GGA functional[25] reduced the error in predicted formation energies of strongly-bound and weakly-bound materials by 50% and 15%, respectively, compared to the PBEsol[26] GGA functional, while simultaneously exhibiting reliable convergence[19]. The original SCAN functional[27] on which it is based has also been shown to predict volumes, lattice constants, and ground-state structures of many materials more accurately than PBE[23,28–32].

Carrying out enough higher-fidelity calculations to comprehensively cover technologically-relevant chemical spaces, as is required for the construction of compositional phase diagrams and the discovery of new structure–property relationships, could clearly benefit the many materials discovery efforts that depend on such data. However, replacing all of the existing lower-fidelity (GGA) calculations in large materials databases with higher-fidelity (e.g., meta-GGA) calculations would consume an enormous amount of energy and computing time, since SCAN and r^{2}SCAN, for example, have 4–5× the computational cost of PBE[19,23,33]. Even if resources were unlimited, there is likely to be little benefit in recomputing materials that are highly unstable (i.e., far from the convex energy hull), since predicting (meta)stability is often of primary importance. Furthermore, meta-GGA calculations will not improve formation energy predictions to an equal extent for all materials. For example, SCAN has been shown to be slightly *less* accurate than PBE in predicting the formation energies of weakly-bound materials (e.g., intermetallics[23]), and r^{2}SCAN improves the predictions for these materials to a much lesser extent than for strongly-bound materials[19].

Therefore, instead of recomputing materials en masse, higher-fidelity calculations should be targeted at the materials for which

[1]Department of Materials Science and Engineering, University of California Berkeley, Berkeley, CA 94720, USA. [2]Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [3]Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [4]Miller Institute for Basic Research in Science, University of California, Berkeley, CA 94720, USA. [5]University of California San Diego, La Jolla, CA 92093, USA. [6]Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. ✉email: kapersson@lbl.gov

they are likely to improve the accuracy of the phase diagrams the most (e.g., strongly-bound materials, materials close to the convex energy hull). Adopting this strategy will economize future use of resources and preserve the massive investment embodied in existing lower-fidelity calculations. Notably, however, such an approach will require phase stability predictions to be based on a mixture of formation energies computed at different levels of theory. The most straightforward way to systematically improve upon GGA phase diagrams in a high-throughput manner, which we refer to here as "naive mixing", is simply to build each phase diagram using formation energies from lower-level calculations, and then replace them with higher-level calculations whenever they are available. However, as we will show, constructing mixed phase diagrams this way can result in severe distortions to the shape of the convex energy hull and dramatically worsen phase stability predictions.

As an alternative to naive mixing, we propose in this work a scheme to construct phase diagrams that mix calculations from different density functionals comprising a lower-fidelity, higher coverage, and a higher-fidelity, lower coverage set of calculations (here, PBE(+U) and r^{2}SCAN) with minimal risk of distortion. By defining the reference state at each point in compositional space as the ground-state PBE(+U) structure, we build a framework in which energies from any two functionals can be mixed in a robust and scalable manner that preserves the shape of the convex hull. After presenting our mixing scheme, we assess how a transition from PBE(+U) to r^{2}SCAN affects predicted polymorph stability and energy above hull by analyzing a set of ~33,900 r^{2}SCAN calculations and discuss strategies for prioritizing r^{2}SCAN calculations such that the mixed phase diagram most closely approximates the full r^{2}SCAN phase diagram. We conclude by using our mixing scheme to analyze solid and aqueous phase stability in two example systems. The mixed phase diagrams presented in this work, along with the 33,000+ new r^{2}SCAN calculations, are made publicly available in the Materials Project database[1] to increase the accuracy of future computational material science efforts.

THEORETICAL FRAMEWORK FOR MIXING FORMATION ENERGIES FROM DIFFERENT FUNCTIONALS

Mixing rules

The computed energy of formation for a material, ΔH_f, is defined with respect to the elements by[10,20,22,34]

$$\Delta H_f^{298\mathrm{K,DFT}} \approx E^{0\mathrm{K,DFT}} - \sum_{\mathrm{el}} n_{\mathrm{el}} E_{\mathrm{el}}^{0\mathrm{K,DFT}} \quad (1)$$

where $E^{0\mathrm{K,DFT}}$ is the total electronic energy computed from DFT at 0 K, subscript 'el' represents each of the constituent elements in the material, and n are stoichiometric coefficients. We note that E can include empirical corrections and that this formulation assumes that differences in finite-temperature enthalpy between materials are negligible. Electronic energies E are not intrinsically meaningful, and their energy scales differ substantially among functionals for the same material. However, the differences in electronic energy among materials and elements, and hence the value of ΔH_f, define a consistent, physically-meaningful quantity that can be compared among different levels of theory.

In this manuscript, we consider mixing PBE(+U) and r^{2}SCAN calculations; although the scheme we present here can be used to mix energies from any two functionals. Note that we use 'PBE(+U)' to refer to the mixture of empirically-corrected PBE and PBE+U calculations that currently populate the Materials Project Database. Specifically, the Materials Project uses PBE for all materials except those containing Co, Cr, Fe, Mn, Mo, Ni, V, and W, which are calculated with a Hubbard U value (See https://docs.materialsproject.org/methodology/gga-plus-u/). These calculations are combined using the mixing scheme of Jain et al.[10] to yield a consistent set of formation energies. In this work, we use this combination of adjusted PBE and PBE+U GGA calculations as our high-coverage, low-fidelity set of calculations, while 'r^{2}SCAN' denotes unadjusted meta-GGA energies that comprise the low-coverage, higher-fidelity calculations. Additional details regarding the computational methods for each calculation type are provided in the Methods section.

As noted in the Introduction, the most straightforward approach to constructing mixed r^{2}SCAN/PBE(+U) phase diagrams is 'naive mixing', where we simply replace PBE(+U) formation energies with r^{2}SCAN formation energies whenever r^{2}SCAN calculations are available, while using PBE(+U) formation energies everywhere else. There are two drawbacks to naive mixing. First, in chemical systems where r^{2}SCAN predicts significantly smaller or larger formation energies than PBE(+U) for most compounds, inserting a single r^{2}SCAN formation energy onto a PBE(+U) phase diagram can either cause that single phase to move off the hull (if $\Delta H_f^{r^2\mathrm{SCAN}} \gg \Delta H_f^{\mathrm{PBE}(+U)}$), or cause adjacent phases to disappear from the hull (if $\Delta H_f^{r^2\mathrm{SCAN}} \ll \Delta H_f^{\mathrm{PBE}(+U)}$), when neither would occur on a full r^{2}SCAN phase diagram. Second, in many cases, r^{2}SCAN stabilizes a different ground-state structure for the elements than PBE(+U) (see Supplementary Note 2). Naive mixing of formation energies in chemical systems containing one of these elements is not rigorously consistent, because the formation energies are being referenced to different structures.

To circumvent these issues, we build our mixing scheme by considering all electronic energies to be the sum of a reference energy and a relative energy. We define the reference energy, E_{ref}, for each functional as *the electronic energy of the PBE(+U) ground-state structure at each point in compositional space*. The energy of any material in either functional may then be expressed as a difference, ΔE_{ref}, relative to the corresponding reference energy. Formation energies are calculated in the usual manner by subtracting the electronic energies of elemental endpoints in each respective functional. Note, however, that ΔE_{ref} is calculated directly from the difference in polymorph energies, and hence does not depend on the energies of the elemental endpoints. Our mixing scheme is similar in spirit to the previous GGA/GGA+U mixing scheme as well as the combined computational–experimental Pourbaix diagrams of the Materials Project[10,35], but extends these approaches to be applicable to any two functionals without relying on pre-fitted energy correction parameters.

Using this framework, we propose two 'mixing rules' that define our scheme for constructing mixed r^{2}SCAN/PBE(+U) phase diagrams. These mixing rules are summarized schematically in Fig. 1 and briefly elaborated below. In sections that follow, we will illustrate each rule with an example.

1. Beginning with a PBE(+U) convex energy hull, replace PBE(+U) energies with r^{2}SCAN energies by adding their ΔE_{ref} to the corresponding PBE(+U) reference energy.
2. Construct the convex energy hull using $\Delta H_f^{r^2SCAN}$ only when there are r^{2}SCAN calculations corresponding to every reference structure (every PBE(+U) stable structure). In this case, add any missing PBE(+U) materials by adding their ΔE_{ref} to the corresponding r^{2}SCAN reference energy.

Rule #1 provides a means to introduce r^{2}SCAN energies onto PBE(+U) phase diagrams when r^{2}SCAN calculations are only available for one or a few compositions. In Fig. 1, polymorph A represents the reference structure (PBE(+U) ground state). Since the reference structure is, by definition, on the PBE(+U) hull, the electronic energy of the r^{2}SCAN relaxed structure corresponding to this reference structure (as determined by the PYMATGEN[36] `StructureMatcher` algorithm) is adjusted to match $E_{\mathrm{ref}}^{\mathrm{PBE}(+U)}$, thereby guaranteeing it will be on the PBE(+U) hull with a

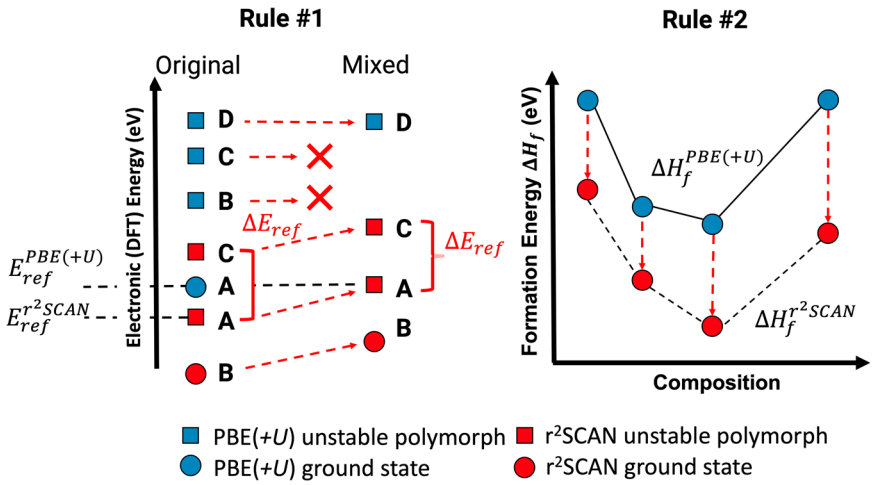

Fig. 1 Rules for mixing PBE(+*U*) (blue) and r²SCAN (red) energies onto a single phase diagram. Left r²SCAN energies can be placed onto a PBE(+*U*) hull by referencing them to the r²SCAN energy of the PBE(+*U*) ground state via ΔE_{ref}. In the diagram, A, B, C, and D represent different polymorphs at a single composition, and polymorph A is the PBE(+*U*) ground state. $E_{ref}^{PBE(+U)}$ and $E_{ref}^{r^2SCAN}$ are the electronic (DFT) energies of polymorph A in the two functionals. PBE(+*U*) energies of polymorphs that also exist in r²SCAN (e.g., polymorphs B and C) are removed. Right The convex hull is built with r²SCAN formation energies only when there are r²SCAN calculations for every PBE(+*U*) ground state.

formation energy equal to $\Delta H_f^{PBE(+U)}$. Electronic energies of Polymorphs B and C, calculated in r²SCAN, are adjusted to maintain their energy difference with respect to the reference structure, ΔE_{ref}. For example, if polymorph C is 10 meV per atom higher in energy than polymorph A in r²SCAN, its energy would be adjusted to $E_{ref}^{PBE(+U)}$ +10 meV per atom. It is also possible for this method to place a polymorph below the PBE(+*U*) hull. Polymorph B is unstable with respect to the reference structure in PBE(+*U*) but is lower in energy than the reference structure in r²SCAN. If it were 10 meV per atom lower than the reference structure in r²SCAN, its energy would be adjusted to $E_{ref}^{PBE(+U)}$ −10 meV per atom, placing it below the original (PBE(+*U*)) hull and slightly changing its shape. Finally, polymorphs that do not have a r²SCAN energy (such as polymorph D) maintain their energy with respect to the PBE(+*U*) hull.

When r²SCAN calculations become available for every reference state, then, by Rule #2, the convex hull is computed directly with r²SCAN formation energies ($\Delta H_f^{r^2SCAN}$). Now, any unstable PBE(+*U*) phases that have not been calculated in r²SCAN can be added to the diagram by adding their ΔE_{ref} to the $\Delta H_f^{r^2SCAN}$ of the corresponding reference structure. This preserves the energy differences between the unstable PBE(+*U*) phases and the corresponding PBE(+*U*) ground state, whose formation energy has now been replaced with an r²SCAN value. In other words, we invert Rule #1 so that r²SCAN structures become the reference structures.

Note that a central assumption of our mixing scheme is that r²SCAN energies are always preferable to PBE(+*U*) energies. This assumption is well-justified by the generally superior accuracy of SCAN and r²SCAN formation energies reported in many studies[19,23,28–32]. In general, the application of our mixing scheme should be restricted to pairs of functionals where one has an a priori reason to prefer one energy over another. The scheme should also only be used when the lower-fidelity calculations make reasonably accurate predictions about the size and shape of the convex energy hull. If this is not the case, the lower-fidelity phase diagram will not be accurate enough to use in the first place and there is no need for a mixing scheme. Finally, we note that in principle it is possible to use our framework to mix energies from more than two functionals, provided that reference energies are available within each functional and that a clear hierarchy can be established among them.

Mixed diagrams for relative polymorph stability (Rule #1)

We illustrate the motivation behind Rules #1 and 2 using the Sn–Br phase diagram, which is shown in Fig. 2. In general, when constructing phase diagrams we seek to determine 1) the shape of the convex energy hull (i.e., stable compositions and their formation energies), and 2) the stable polymorph at each composition. Figure 2a and b compare the Sn–Br phase diagram with the formation energy of all phases calculated in PBE(+*U*) and r²SCAN, respectively, and show that the accuracy of both aspects is improved by r²SCAN. PBE(+*U*) incorrectly predicts the ground-state polymorph of $SnBr_2$ as rocksalt (spacegroup $P\bar{3}m1$) and overpredicts the magnitude of ΔH_f as −1.136 eV per atom, whereas the experimental value is estimated at −0.84 to −0.92 eV per atom[37] (indicated by the shaded band in Fig. 2). By contrast, r²SCAN correctly predicts the $SnBr_2$ ground state as *Pnma*[29] and makes a substantially more accurate prediction of its formation energy (−0.833 eV per atom).

As we have discussed, it is not always feasible to recompute an entire chemical system using r²SCAN (as we have done to construct Fig. 2b). When improving predictions of polymorph stability is a primary research objective, it makes sense to prioritize r²SCAN calculations for all known polymorphs at the composition of interest. However, if we were to apply this strategy to $SnBr_2$ and replace all PBE(+*U*) formation energies of $SnBr_2$ polymorphs with r²SCAN energies using naive mixing (Fig. 2c), $SnBr_2$ would no longer be predicted as stable. This occurs because the entire hull is shallower (smaller magnitude of ΔH_f) in r²SCAN than in PBE(+*U*), and hence using an r²SCAN formation energy for $SnBr_2$ causes it to move off the hull. Instead, we must apply Rule #1 to make the r²SCAN energies compatible with the PBE(+*U*) hull. We do so by positioning r²SCAN formation energies relative to the PBE(+*U*) ground state polymorph ($P\bar{3}m1$), as shown in Fig. 2d. Because we maintain the energy differences relative to this reference energy, the correct polymorph is now stabilized.

Compared to naive mixing of formation energies, applying Rule #1 preserves the overall shape of the PBE(+*U*) convex hull while enabling improvement in phase stability predictions using as few as two r²SCAN calculations (one for the polymorph of interest and one for the reference structure). However, because r²SCAN stabilizes the *Pnma* polymorph instead of the $P\bar{3}m1$ polymorph stabilized by PBE(+*U*), ΔH_f is lowered (and made less accurate) by 37 meV per atom, which is the difference in energy between the $P\bar{3}m1$ and *Pnma* polymorphs in r²SCAN. Hence, although use of

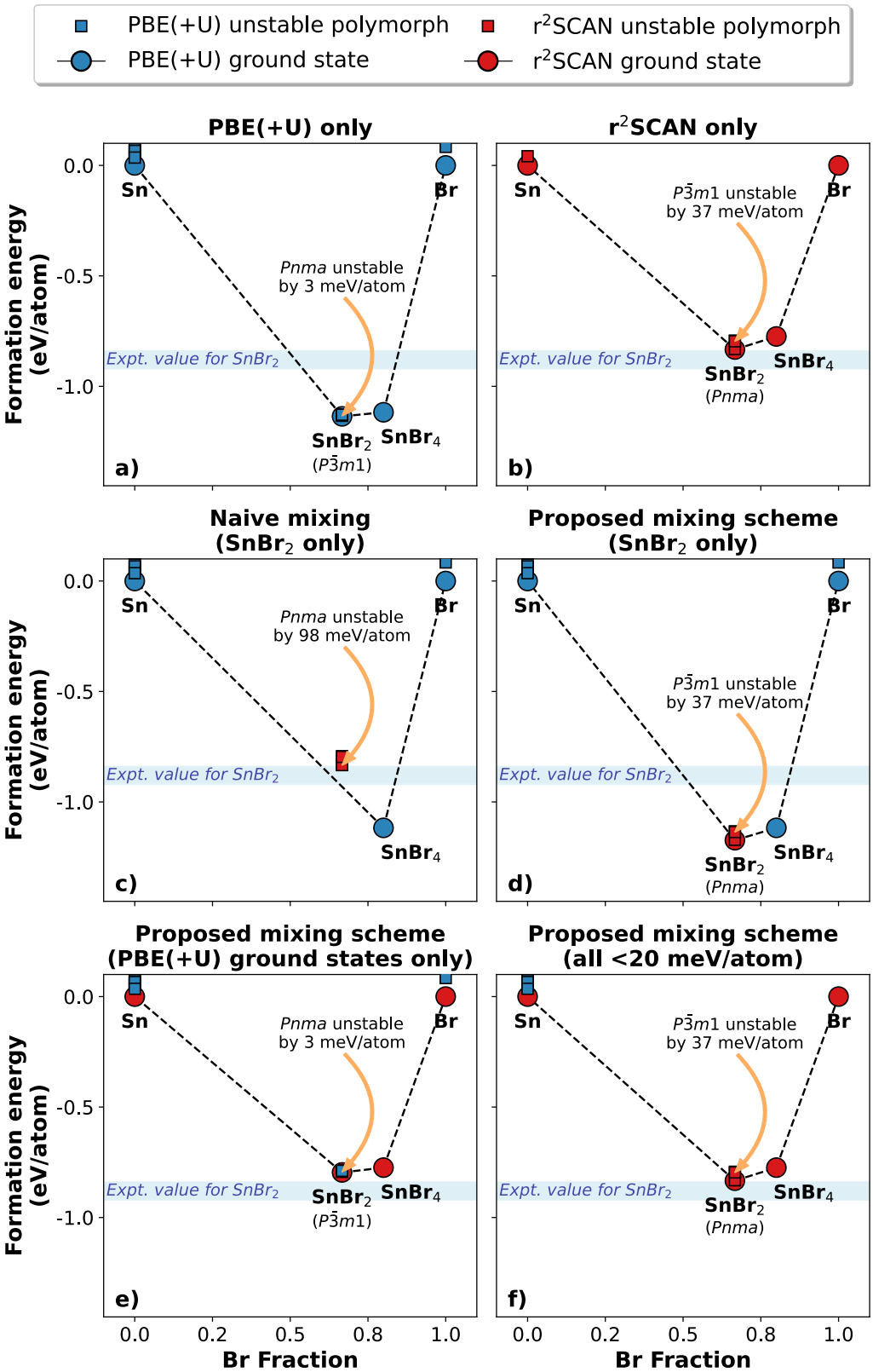

Fig. 2 Sn–Br phase diagram constructed using different mixing strategies for PBE(+*U*) and r²SCAN calculations. a PBE(+*U*) only; **b** r²SCAN only; **c** naive mixing of formation energies, where all $SnBr_2$ polymorphs are computed in r²SCAN and all other materials are in PBE; **d** the same set of energies as **c**, but employing our mixing scheme (Rule #2); **e** r²SCAN for all reference states (i.e., PBE(+*U*) ground states) and PBE(+*U*) for all other materials; **f** r²SCAN for all materials within 20 meV per atom of the PBE(+*U*) convex hull. The numerical value in parentheses indicates the energy above hull of the experimental ground state *Pnma* polymorph of $SnBr_2$. The shaded blue regions represent the estimated range of experimental formation energies for $SnBr_2$[37]. Tabulated r²SCAN and PBE(+*U*) energies for all materials are available as Supplementary Data (see Supplementary Note 4).

Rule #1 for study of a single composition may yield more accurate relative polymorph energies, it carries the risk of making the magnitude of the formation energy slightly less accurate compared to a full r^{2}SCAN phase diagram.

Mixed diagrams for formation energy (Rule #2)

When identifying stable compositions or predicting accurate formation energies is the primary research objective, it makes sense to prioritize recomputing all PBE(+U) ground states in r^{2}SCAN, as shown in Fig. 2e. Applying Rule #2, this strategy allows the entire convex hull to be constructed using $\Delta H_f^{r^2SCAN}$. Unstable PBE(+U) polymorphs are then positioned relative to the corresponding reference structures. Several unstable polymorphs of Sn that were mixed in this manner are visible in Fig. 2e.

In this chemical system, recomputing the hull when only the PBE(+U) ground states have been calculated in r^{2}SCAN will still not recover the exact r^{2}SCAN formation energy for $SnBr_2$, because PBE(+U) stabilizes the incorrect ground state, and Rule #2 treats this incorrect ground state as the reference energy. Hence, the formation energy of $SnBr_2$ predicted by the mixed phase diagram in Fig. 2e is too high by 37 meV per atom (the difference in energy between the $P\bar{3}m1$ and *Pnma* polymorphs in r^{2}SCAN).

Because there is no way to know a priori whether r^{2}SCAN will stabilize a different ground state than PBE(+U), a more robust strategy is to compute all polymorphs within some tolerance of the PBE(+U) convex energy hull. Computing both ground states and slightly metastable polymorphs with r^{2}SCAN makes it more likely that the shape of the convex energy hull in the mixed phase diagram will be identical to that in a full r^{2}SCAN diagram. We apply this strategy in Fig. 2f, in which we mix r^{2}SCAN energies for all materials within 20 meV per atom of the PBE(+U) hull, and use PBE(+U) energies for all other materials. The value of 20 meV per atom is motivated by analysis presented later (see Section "Definition of materials 'close to the hull'") indicating that materials with higher $\Delta E_{hull}^{PBE(+U)}$ are rarely stabilized by r^{2}SCAN. With this strategy, the shape of the resulting energy hull (Fig. 2f) exactly matches that of the pure r^{2}SCAN hull (Fig. 2b).

Comparing Fig. 2a, c, e illustrates the importance of Rule #2 when mixing r^{2}SCAN and PBE(+U) calculations. Formation energies cannot be naively mixed without carrying a substantial risk of over- or under-stabilizing certain compositions. The convex energy hull must remain in PBE(+U) until there are r^{2}SCAN calculation corresponding to every reference state. Even in that case, it is preferable to include slightly unstable polymorphs in order to achieve better accuracy in cases where r^{2}SCAN stabilizes different polymorphs.

PRACTICAL CONSIDERATIONS

Mixed diagrams for ternary and higher systems

Ternary or higher-dimensional chemical spaces present special challenges for mixing energies between functionals, because strict application of mixing rules #1 and #2 can introduce inconsistencies between the full phase diagrams and those of constituent subsystems. For example, consider a case in which all of the PBE(+U) ground states in chemical system A–B are computed in r^{2}SCAN. According to Rule #2, the binary A–B phase diagram would be constructed using $\Delta H_f^{r^2SCAN}$. This may result in different formation energies and/or predicted stable phases than the PBE(+U) phase diagram, as illustrated previously for the Sn–Br system. Now suppose that we wish to construct a ternary phase diagram for the A–B–C system, in which there are multiple ternary ground states that have not been computed in r^{2}SCAN. Since the A–B–C system does not satisfy the requirements for Rule #2, we would construct this ternary phase diagram using $\Delta H_f^{PBE(+U)}$. This could result in the ternary A–B–C diagram predicting different formation energies and/or stable phases in the A–B subsystem than the binary A–B diagram. Such inconsistency may be problematic depending on the use case. Note that if all reference energies in the full A–B–C system have been recomputed in r^{2}SCAN, then strict application of the mixing rules will not result in any inconsistencies. However, due to the much larger number of ternary and higher materials (compared to binaries), it becomes progressively more difficult to recompute all the reference energies needed to apply Rule #2 as the size of the chemical system increases.

In cases where consistency between lower- and higher-dimension phase diagrams is essential, one may apply the mixing rules individually to each chemical subsystem, in order of increasing dimensionality. To continue the example above, Rule #1 and Rule #2 would be applied individually to each of the A–B, B–C, and A–C chemical systems. The ternary phase diagram would then be constructed by combining these pre-adjusted binary formation energies with PBE(+U) formation energies. An example of such a diagram is presented in Supplementary Fig. 9.

Applying the mixing scheme to binary subsystems before treating the ternary system amounts to a modified form of naive mixing because it involves directly combining formation energies obtained from PBE(+U) (for ternaries) with those calculated with r^{2}SCAN for binaries, without considering whether r^{2}SCAN energies are available for all ternary ground states. As such, mixed phase diagrams for high-dimensional chemical systems that are constructed in this manner should be used sparingly and interpreted with care. However, due to the inherently larger number of phases involved in higher-dimensional systems, we expect this modified form of naive mixing to be less likely to cause severe distortions of the hull compared to binary systems.

To test this hypothesis, we compared ternary phase stability predictions from ~6000 ternary phase diagrams computed in PBE(+U) to mixed versions constructed using $\Delta H_f^{r^2SCAN}$ values for all binary subsystems and $\Delta H_f^{PBE(+U)}$ for all ternary materials (see Supplementary Figs. 8 and 9). We evaluated how frequently these 'edged' diagrams either 1) destabilized a known experimental ternary phase (i.e., a phase reported in the Inorganic Crystal Structure Database[38]) that was stable in pure PBE(+U) or 2) stabilized a known experimental ternary phase that was unstable in pure PBE(+U). For the majority of chemical systems (83%), experimental ternary materials predicted stable by the pure PBE(+U) diagram remained so in the mixed phase diagram, while in another 14% of cases exactly one material was destabilized. Similarly, for 94% of chemical systems, experimental materials predicted unstable by the pure PBE(+U) diagram remain so in the mixed diagram, while for 6% of chemical systems exactly one unstable experimental material was stabilized (see Supplementary Fig. 8). Thus, although employing modified naive mixing (i.e., 'edged' phase diagrams) to achieve consistency between lower- and higher-dimensional phase diagrams carries a modest risk of destabilizing known experimental phases for some chemical systems, there are many other cases in which the mixed diagrams stabilize experimental phases that pure PBE(+U) does not. Altogether, these results suggest that modified naive mixing is unlikely to *severely* distort phase stability predictions for ternary or higher dimensional chemical systems.

Definition of materials 'close to the hull'

In the previous section "Theoretical framework for mixing formation energies from different functionals" we observed that it is preferable to recompute not just PBE(+U) ground states, but also materials close to the convex energy hull in order to ensure that the mixed energy hull has the correct shape (compare Fig. 2e, f). This begs the question of how to define 'close to the hull'. More specifically, we can rephrase the question as 'how likely is r^{2}SCAN to stabilize a material that is X meV per atom above the hull in PBE(+U)?' For example, a material that is 500 meV per atom above

hull in PBE(+*U*) will almost certainly not become stable in r^{2}SCAN, but a material that is unstable by 3 meV per atom could (as was the case with *Pnma* $SnBr_2$ in Fig. 2). Determining an appropriate threshold is necessary to properly target r^{2}SCAN calculations.

To inform this question, in Supplementary Note 1 we evaluate the extent to which r^{2}SCAN changes the energy above hull of unstable polymorphs. Examining approximately 7300 unstable materials with a PBE(+*U*) energy above hull of 50 meV per atom or less, we find that in 95% of cases, the energy above hull either *increases* or decreases by no more than 19 meV per atom (see Supplementary Fig. 1). This means that materials more than 19 meV per atom above the PBE(+*U*) hull would only be stabilized by r^{2}SCAN in rare cases. Hence, we adopt a threshold of 20 meV per atom as our definition of 'close to the hull' for purposes of prioritizing calculations.

By way of comparison, we note that among 16 systems identified by Yang et al.[29] in which SCAN stabilized the correct ground state and PBE(+*U*) did not, the energies above hull of the experimental ground states in PBE(+*U*) ranged from 2 to 50 meV per atom. Another study showed that SCAN mispredicted the ground states of TiO_2 and FeS_2, with misprediction on the order of 50 meV per atom as well[39]. Thus, although based on analysis of a large set of materials, our selection of 20 meV per atom as a safe threshold is not guaranteed to capture the r^{2}SCAN ground state polymorph in every case. A higher threshold could certainly be chosen if greater confidence in capturing the correct ground states is required.

Failures of structure matching

In the Section "Mixing rules" we established the need to obtain r^{2}SCAN energies of PBE(+*U*) ground states, which serve as reference energies for constructing mixed phase diagrams. To obtain the most accurate r^{2}SCAN energies, we generally perform r^{2}SCAN structure optimizations calculations and then use the PYMATGEN[36] `StructureMatcher` algorithm to determine whether the r^{2}SCAN-relaxed structure is the same (within tolerances) as the PBE(+*U*) starting structure.

In the vast majority of cases, the r^{2}SCAN-relaxed structure and the PBE(+*U*) starting structure match, allowing us to use the r^{2}SCAN energy as reference energy. However, in selected cases (some 1% of all materials we have computed thus far), r^{2}SCAN will optimize to a structure that is no longer considered equivalent to the starting structure. This is especially common for the crystal structures of diatomic molecules (e.g., H_2, Cl_2, O_2) in which the different treatment of short- and medium-range interactions by r^{2}SCAN compared to PBE is particularly significant.

We address this issue in two ways. In some cases, manual inspection of the structures allows us to establish that they represent the same material, and hence that the r^{2}SCAN energy can be used as a reference energy. However, manual inspection is not feasible for high-throughput work. Instead, we perform single-point calculations for any materials in which the r^{2}SCAN-relaxed structure no longer matches the input structure. The r^{2}SCAN single point calculation is guaranteed to match the corresponding PBE(+*U*)-optimized structure and provides an r^{2}SCAN energy that can serve as a reference energy. Meanwhile, an r^{2}SCAN optimization of the same structure (which may no longer be the same according to the `StructureMatcher`) is guaranteed to have a similar or lower energy than the single point and will be added to the hull at the correct position by application of Rule #1. Performing r^{2}SCAN single points also provides a means of obtaining reference energies for large structures that would be impractical to optimize in r^{2}SCAN within reasonable computational limits (see next Section).

Prioritizing r^{2}SCAN calculations for maximum benefit

We conclude our discussion of practical considerations by considering the best strategy for prioritizing r^{2}SCAN calculations, given that computational resources are limited and that its cost is still ~5× that of PBE[19].

We can define several levels of "calculation coverage" (meaning, subsets of materials that have all been recomputed with r^{2}SCAN, Fig. 3) based on the mixing rules we have established. In order to apply Mixing Rule #1, at least two r^{2}SCAN optimizations at a single composition are needed: one for the PBE(+*U*) ground state and one for another polymorph. To apply Rule #2, we require r^{2}SCAN energies for every PBE(+*U*) ground state or (ideally) every PBE(+*U*) material within 20 meV per atom of the hull. These energies are preferably obtained from structure optimizations, although as elaborated below, single-point calculations can be used in select cases, with the risk of a slightly less accurate hull shape. The pinnacle of calculation coverage (which may have less value than its computational cost, as noted in the Introduction) is full recomputation of all materials using r^{2}SCAN.

With a goal of achieving second- or third-level coverage, we can identify several strategies for prioritizing which materials to calculate in order to maximize the benefits of the mixing scheme for formation energy prediction. To do so, we classify materials close to the hull as (1) strongly- or weakly-bound and (2) small or large.

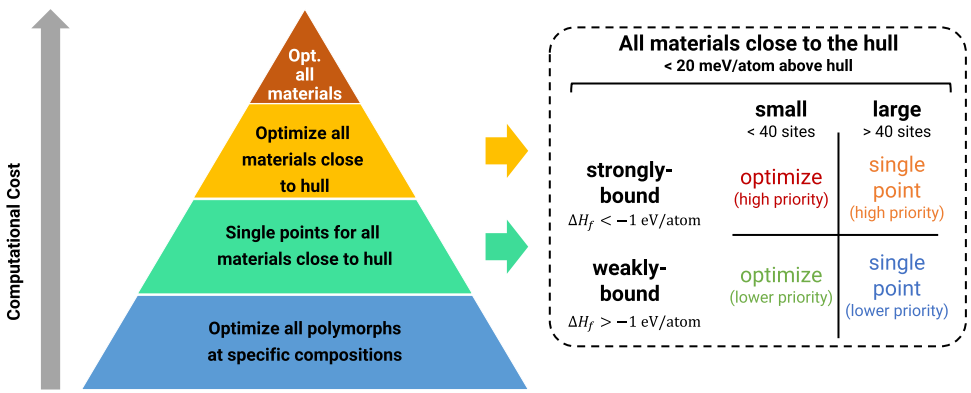

Fig. 3 Strategies for prioritizing higher level calculations in large materials databases. Left different stages of calculation coverage. The first level enables application of Mixing Rule #1, while the second and third levels facilitate application of Rule #2. The fourth level (re-optimizing all materials with higher-level calculations) is not necessary. Right methods used by the Materials Project to pursue second- and third-level coverage of r^{2}SCAN calculations. Optimizations of small, strongly-bound materials have the highest priority. single point calculations are used for large structures. Weakly-bound materials have lower priority. See Section "Prioritizing r^{2}SCAN calculations for maximum benefit" for further details.

Previous studies[19,23] established that SCAN and r^{2}SCAN predict substantially more accurate formation energies than PBE or PBEsol for "strongly-bound" materials, i.e., materials whose PBE(+*U*)-predicted formation energy is lower than −1 eV per atom. The improvement in accuracy for weakly-bound materials is more modest. As such, creating mixed phase diagrams for strongly bound systems is likely to improve overall accuracy the most, and hence we assign higher priority to strongly-bound materials.

With respect to size, experience indicates that optimizations of large structures (e.g., larger than ~40 sites) with r^{2}SCAN will often exceed typical maximum wall time limits at supercomputing centers (e.g., 48 hr). This is not to say optimization is impossible; rather, in a high-throughput computing context, it does not usually make sense to invest an excessive amount of computing nodes or wall time into a single material. As such, in these cases, we may perform single-point calculations in order to obtain reference energy (albeit a less accurate one) so that Rule #2 can be applied. Fully-optimized structures can be obtained as computational resources allow, and added into the mixed phase diagrams according to Rule #1. We emphasize that single-point calculations are only used in very limited cases as a practical necessity, and all r^{2}SCAN energies presented herein are based on full optimizations rather than single points. In general, a formation energy derived from an r^{2}SCAN single point energy cannot be assumed to be more accurate than one derived from a PBE(+*U*) relaxation; hence it is not advisable to rely exclusively on such calculations in order to apply the mixing scheme. However, phase diagrams of technologically-relevant chemical systems often contain many small structures and only a few large ones that are impractical to optimize with r^{2}SCAN. In such cases, the use of single-point calculations for the large structures makes it possible to apply Rule #2 and build the convex energy hull using r^{2}SCAN energies, which presumably offers an overall improvement in predictive accuracy for the chemical system that outweighs the possible loss of accuracy in the formation energy of the few large structures.

EXAMPLES

Application to a metastable ternary nitride system

As a practical example of our complete mixing scheme, we use it to investigate compound metastability in the ternary Zn–Sb–N system. Nitrides remain relatively unexplored compared to other chemical spaces, even though they exhibit the largest range of thermodynamically-accessible metastable states among inorganic materials[40,41], which is thought to be a consequence of the large cohesive energy of metal-nitrogen bonds that kinetically traps metastable structures[42]. Compared to stable nitrides, metastable nitrides are more likely to contain metal cations in high oxidation states, which imparts unique semiconducting properties that make these materials interesting for electronic and photovoltaic applications, among others[30,41,42].

Metastable nitrides are relatively rare in nature and difficult to synthesize experimentally due to the high stability of molecular N_2. However, the use of reactive nitrogen precursors such as ammonia, azide compounds, or plasma-cracked atomic N allows nitrogen chemical potentials of up to +1 eV perN above the convex energy hull to be reached in laboratory synthesis[30,42,43]. Recent experimental studies[44–46] have reported synthesis of several metastable nitrides (Cu_3N, Sn_3N_4, and Ti-alloyed Sn_3N_4). Ternary Wurtzite-based nitrides, such as $MgSnN_2$, $ZnSnN_2$, $ZnGeN_2$, and $ZnSiN_2$, have received specific attention recently as potential alternatives to III-V semiconductors[43,44]. Computational screening studies[30,42] recently predicted three new metastable ternary phases ($ZnSb_2N_4$, Zn_2SbN_3, and Zn_3SbN_3) and a new metastable binary phase (SbN) in the Zn–Sb–N chemical space. Zn_2SbN_3, the first Sb-based nitride semiconductor ever reported, was experimentally realized[42,43,47] and exhibited promising electronic properties for photovoltaic and water-splitting applications. SbN was predicted to be relatively close to the metastability limit (requiring +0.8 eV perN to stabilize)[30] and is the subject of ongoing investigations as another potential Sb-based nitride semiconductor.

Given the diverse bonding characteristics of nitrogen compounds, computational predictions of metastability can be particularly sensitive to the choice of functional and energy correction scheme (e.g., PBE vs. PBE+*U* vs. r^{2}SCAN). For example, Sun et al.[30] showed that for many binary nitride systems, PBE overstabilizes the nitrogen-rich region of the convex energy hull, while PBE+U overstabilized the nitrogen-poor region. SCAN was found to predict formation enthlapies with good accuracy across both portions of the hull[30]. To expand on this previous work and inform future high-throughput screening studies, we evaluate how the use of a mixed r^{2}SCAN/PBE(+*U*) ternary phase diagram would affect these predictions.

Figure 4a and b show the phase diagrams computed entirely using PBE(+*U*) and r^{2}SCAN calculations, respectively. Both the pure PBE(+*U*) and pure r^{2}SCAN phase diagrams predict that Zn_2SbN_3 and SbN are metastable, consistent with the previous studies. However, the energy above hull of both metastable compositions of interest is higher in r^{2}SCAN. r^{2}SCAN predicts 30 meV per atom above hull for Zn_2SbN_3 (vs. 20 meV per atom in PBE), reflecting the overstabilization of the N-rich region of the phase diagram in PBE(+*U*) noted by Sun et al.[30] r^{2}SCAN predicts 260 meV per atom above hull for SbN (vs. 172 meV per atom in PBE), but in the r^{2}SCAN diagram this material falls within the metastable synthesizability limit, consistent with experimental reports[30]. The other notable differences between the PBE(+*U*) and r^{2}SCAN phase diagrams are that Zn_3SbN_3, which has not been synthesized to the best of our knowledge and is not in the ICSD, is predicted stable by PBE(+*U*) but unstable by r^{2}SCAN, while $ZnSb_2N_4$ is unstable in PBE(+*U*) but metastable in r^{2}SCAN. The stable/metastable/unstable classification for all other compositions is the same in both diagrams.

Moving to mixed phase diagrams, we now consider a situation in which only the elements and binary compositions have been computed with r^{2}SCAN, while all ternary phases remain in PBE(+*U*). We compare two methods of constructing this mixed ternary phase diagram in Fig. 4c, d. In Fig. 4c, we apply Mixing Rule #1 and #2 strictly (i.e., considering the entire phase diagram at once). In this scenario, because we do not have a r^{2}SCAN calculation for the PBE(+*U*) reference structure Zn_3SbN_3, *the hull is still calculated using PBE(+U) energies* (with the exception of polymorphs stablized by r^{2}SCAN, as discussed later). r^{2}SCAN polymorphs for each element or binary composition are placed on this PBE(+*U*) hull by anchoring to the respective reference states according to Rule #1.

Inspection of Fig. 4c shows broad similarity to the pure r^{2}SCAN diagram (Fig. 4b), with a few notable differences. Zn_3SbN_3 is predicated unstable in the pure r^{2}SCAN diagram yet metastable in the mixed diagram. For SbN the reverse is true: this material is predicted metastable in the pure r^{2}SCAN diagram but unstable in this mixed diagram. The convex energy hull in Fig. 4c is constructed with PBE(+*U*) energies, so it is identical to that of Fig. 4a with one significant exception. As noted in Section "Mixed diagrams for relative polymorph stability (Rule #1)", Rule #1 can cause the energy of the convex hull to decrease in cases where r^{2}SCAN stabilizes a different polymorph than PBE(+*U*). In this case, r^{2}SCAN stabilizes a different structure for N_2 (which is a crystalline solid at 0 K) that is 1.8 meV per atom *lower* in energy than the reference energy (see Supplementary Fig. 2), causing it to be placed below the PBE(+*U*) hull and thereby *lowering* the hull energy of the N-rich region in the mixed phase diagram. This causes Zn_3SbN_3, which is predicted stable in the pure PBE(+*U*) phase diagram, to move off the hull by 0.4 meV per atom and

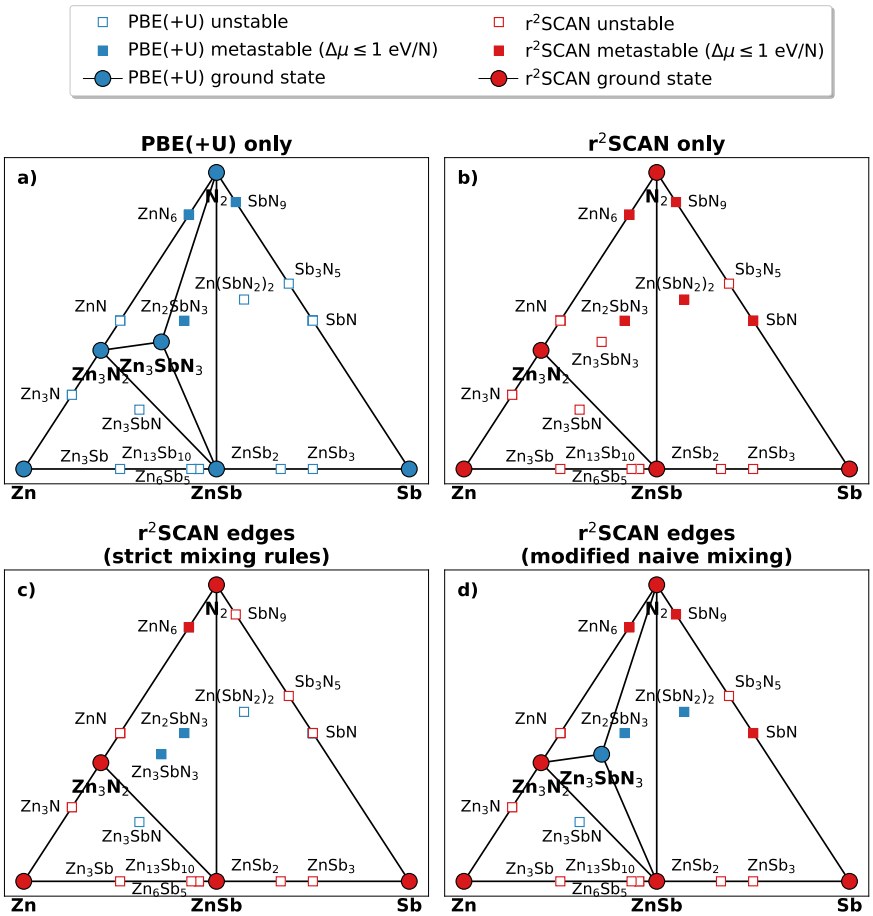

Fig. 4 Zn–Sb–N phase diagrams illustrating different mixing strategies for PBE(+*U*) and r²SCAN calculations. a PBE(+*U*) only; **b** r²SCAN only; **c** strict application of Rule #1 and #2 to calculations comprising r²SCAN energies for all elements and binary phases with PBE(+*U*) energies for all ternary phases (**d**) same set of calculations as (**c**), but using modified naive mixing in which binary hulls are constructed from r²SCAN formation energies. Phases labeled `metastable' are phases that can be stabilized by a +1 eV perN increase in the nitrogen chemical potential, which is achievable in laboratory synthesis[30,42,43]. Tabulated r²SCAN and PBE(+*U*) energies for all materials are available as Supplementary Data (see Supplementary Note 4).

become metastable. The energy above hull for SbN increases by 1 meV per atom, and hence it retains its classification as unstable consistent with the PBE(+*U*) phase diagram.

Figure 4d presents an alternative phase diagram constructed by fully applying Rules #1 and #2 to the binary edges and then adding PBE(+*U*) energies for ternary phases by modified naive mixing, as discussed in the Section "Mixed diagrams for ternary and higher systems". Here, the edges of the diagram are identical to those predicted by a pure r²SCAN diagram because we have full coverage of all PBE(+*U*) ground states and hence Rule #2 applies. In the interior of the diagram, three of the four ternary compositions retain the same stable/unstable/metastable classification they have in the pure r²SCAN diagram, while Zn_3SbN_3 is predicted to be stable in the PBE(+*U*) diagram (whereas it is predicted unstable in the pure r²SCAN diagram).

As noted above in the Section "Mixed diagrams for ternary and higher systems", the modified form of naive mixing employed to construct Fig. 4d is thermodynamically less consistent than strict application of Rule #1 and #2 (Fig. 4c), and should only be invoked when *consistency between binary and higher-dimension phase diagrams is essential*. In this example, where the metastability of phases is of primary interest, it would be advisable to apply the mixing rules strictly to ensure that the entire convex hull is constructed in a consistent manner. Indeed, among the two mixed diagrams, Fig. 4c shows the most consistency with the pure r²SCAN diagram.

Application to aqueous phase stability

Finally, we demonstrate how the mixing scheme presented here can be used to inform aqueous phase stability predictions. The computational Pourbaix diagram formalism of Persson et al.[35,48] generates aqueous stability (pH–pE) diagrams by referencing experimental free energies of dissolved ions to DFT-predicted formation energies derived from solid phase diagrams. As such, the mixing scheme presented here can be applied to the creation of Pourbaix diagrams in addition to solid phase diagrams. SCAN-derived Pourbaix diagrams, for example, were shown to be systematically more accurate for transition metal oxides[49]. However, the large number of stable phases needed to build computational Pourbaix diagrams may preclude calculating entire chemical spaces in SCAN or r²SCAN, motivating the usefulness of our mixing scheme in this context.

We illustrate the mixing scheme on the SeO system, for which the PBE-derived Pourbaix diagram is known to be inaccurate with respect to experiment[49]. Specifically, it predicts a stable SeO_2 phase that is not observed experimentally (Fig. 5a). Creating a computational Pourbaix diagram of this system requires a solid phase diagram of the Se-O–H chemical system, which contains 85 individual materials, according to the Materials Project database. Five of these materials contain >40 sites, and hence could be particularly challenging to recompute in r²SCAN (see Section "Prioritizing r2SCAN calculations for maximum benefit"). Use of the mixing scheme allows us to construct the hull in r²SCAN by

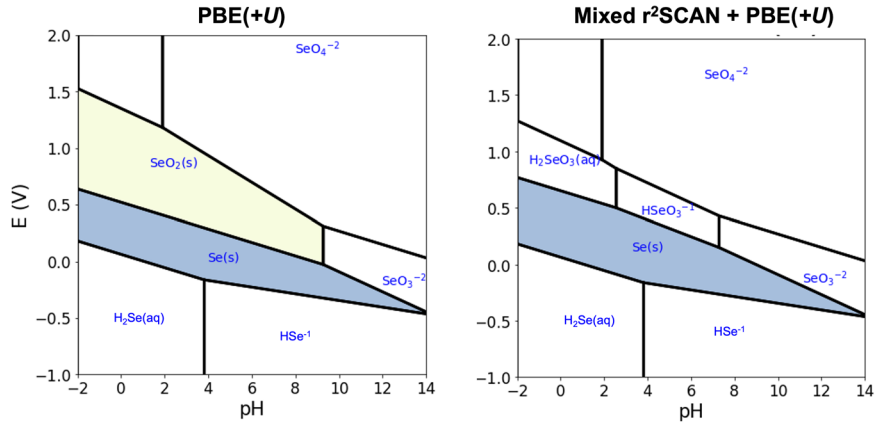

Fig. 5 Application of the mixing scheme to aqueous phase stability. Pourbaix diagram of Se–O constructed using PBE(+*U*) calculations (left) and the mixing scheme presented in this work (right) at a total Se concentration of 10^{-6} mol/L. Shaded regions predict stable solid phases. Tabulated r^{2}SCAN and PBE(+*U*) energies for all materials are available as Supplementary Data (see Supplementary Note 4).

performing calculations only for the ground states (9 materials) while retaining information from PBE(+*U*) about the metastability of other phases.

The resulting Pourbaix diagram built from mixed r^{2}SCAN and PBE(+*U*) energies (Fig. 5b) correctly predicts that the oxide phase SeO_2 is unstable (unlike the pure PBE(+*U*) diagram), in agreement with Pourbaix diagrams presented by Wang et al.[49] that were prepared from both experimental data and pure SCAN calculations. Hence for this system, our mixing scheme has made it possible to leverage a relatively small number of calculations and achieve similar predictive accuracy as full recalculation of all materials with SCAN.

In summary, we have developed a mixing scheme to enable construction of phase diagrams that combine formation energies from any two DFT functionals. Such a capability is important to high-throughput materials screening efforts because it allows a relatively low-coverage, high-fidelity set of calculations (here, r^{2}SCAN) to be used in concert with existing high-coverage, lower-fidelity calculations (here, PBE(+*U*)) to improve the accuracy of phase stability predictions. Our scheme allows mixing r^{2}SCAN and PBE(+*U*) calculations when as few as two r^{2}SCAN calculations (one corresponding to a PBE(+*U*) ground state) are available, and scales smoothly to cases where entire binary, ternary, or higher-dimensional chemical systems are calculated in r^{2}SCAN. We identified specific guidelines that can be used to target limited computational resources towards materials where r^{2}SCAN calculations are likely to improve accuracy the most, and illustrated how the mixing scheme can be applied to solid and aqueous phase stability predictions.

METHODS

We employed the Vienna ab initio Simulation Package (VASP)[50,51] v.6.1.1 in conjunction with v.54 of the projector-augmented wave (PAW) PBE pseudopotentials[52] for all r^{2}SCAN calculations in this work. We employed a two-step high-throughput workflow described in greater detail elsewhere[19], which comprises a structure optimization with PBEsol[26] to generate an initial guess of the charge density, followed by a subsequent structure optimization with r^{2}SCAN[25]. Briefly, these calculations were carried out with a plane-wave energy cutoff of 680 eV, a bandgap-dependent *k*-point density[19,53], and force-convergence tolerance of 0.02 eV per Å (`EDIFFG=−0.02` in VASP), which were developed to achieve formation energy converged to within ~1 meV per atom. The `ACCURATE` precision keyword was enabled in VASP, aspherical contributions to the PAW spheres were included, and the self-consistent field was converged to at least 10^{-5} eV. Spin-polarization was considered for all calculations.

We do not apply any energy corrections to the resulting r^{2}SCAN energies. For context regarding corrections, in Supplementary Note 3 we fit energy corrections to diatomic gases and show that the corrections that would be applied to r^{2}SCAN are substantially smaller than those that have been widely used for PBE(+*U*)[15,16].

PBE calculations were retrieved from the Materials Project REST API[54]. Calculations for transition metal oxides and fluorides contained a Hubbard *U* value and incorporated the GGA/GGA+*U* mixing scheme of Jain et al.[10], in addition to empirical corrections applied to some chemical systems[11]. We refer to these calculations as PBE(+*U*) throughout this work.

DATA AVAILABILITY

All data referenced herein are publicly available in the Materials Project database[54]. At the time of this publication the database contains r^{2}SCAN calculations for ~33,000 materials, corresponding to 77% of all elements, binary, and ternary materials within 20 meV per atom of the PBE(+*U*) convex energy hull.

CODE AVAILABILITY

Our computational workflow has been implemented into the PYMATGEN[36] and ATOMATE[55] packages as of version 2020.1.28 and 0.9.5, respectively, for readers wishing to utilize it in their own work. The mixing scheme described herein is available in the `MaterialsProjectDFTMixingScheme` class in PYMATGEN[36] as of release 2022.1.20.

Received: 3 March 2022; Accepted: 25 August 2022;
Published online: 13 September 2022

REFERENCES

1. Jain, A. et al. The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
2. Curtarolo, S. et al. AFLOWlib.org: A distributed materials properties repository from high-throughput ab initio calculations. *Comput. Mater. Sci.* **58**, 227–235 (2012).
3. Saal, J. E., Kirklin, S., Aykol, M., Meredig, B. & Wolverton, C. Materials design and discovery with high-throughput density functional theory: The open quantum materials database (OQMD). *JOM* **65**, 1501–1509 (2013).
4. Kirklin, S. et al. The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. *npj Comput. Mater.* **1**, 15010 (2015).
5. Choudhary, K. et al. The joint automated repository for various integrated simulations (JARVIS) for data-driven materials design. *npj Comput. Mater.* **6**, 1–13 (2020).

6. Himanen, L., Geurts, A., Foster, A. S. & Rinke, P. Data-driven materials science: Status, challenges, and perspectives. *Adv. Sci.* **6**, 1900808 (2019).
7. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
8. Langreth, D. C. & Perdew, J. P. Theory of nonuniform electronic systems. I. analysis of the gradient approximation and a generalization that works. *Phys. Rev. B* **21**, 5469 (1980).
9. Dudarev, S. L., Botton, G. A., Savrasov, S. Y., Humphreys, C. & Sutton, A. P. Electron-energy-loss spectra and the structural stability of nickel oxide: an LSDA+U study. *Phys. Rev. B* **57**, 1505 (1998).
10. Jain, A. et al. Formation enthalpies by mixing GGA and GGA+U calculations. *Phys. Rev. B.* **84**, 1–10 (2011).
11. Wang, A. et al. A framework for quantifying uncertainty in DFT energy corrections. *Sci. Rep.* **11**, 15496 (2021).
12. Chen, C. et al. A critical review of machine learning of energy materials. *Adv. Energy Mater.* **10**, 1903242 (2020).
13. Mori-Sánchez, P., Cohen, A. J. & Yang, W. Many-electron self-interaction error in approximate density functionals. *J. Chem. Phys.* **125**, 201102 (2006).
14. Perdew, J. P. Climbing the ladder of density functional approximations. *MRS Bull.* **38**, 743–750 (2013).
15. Grindy, S., Meredig, B., Kirklin, S., Saal, J. E. & Wolverton, C. Approaching chemical accuracy with density functional calculations: diatomic energy corrections. *Phys. Rev. B* **87**, 1–8 (2013).
16. Wang, L., Maxisch, T. & Ceder, G. Oxidation energies of transition metal oxides within the GGA+U framework. *Phys. Rev. B* **73**, 1–6 (2006).
17. Anisimov, V. I., Zaanen, J. & Andersen, O. K. Band theory and mott insulators: Hubbard *U* instead of Stoner *I*. *Phys. Rev. B* **44**, 943–954 (1991).
18. Grimme, S., Hansen, A., Brandenburg, J. G. & Bannwarth, C. Dispersion-corrected mean-field electronic structure methods. *Chem. Rev.* **116**, 5105–5154 (2016).
19. Kingsbury, R. et al. Performance comparison of r^2SCAN and SCAN metaGGA density functionals for solid materials via an automated, high-throughput computational workflow. *Phys. Rev. Mater.* **6**, 013801 (2022).
20. Lany, S. Semiconductor thermochemistry in density functional calculations. *Phys. Rev. B.* **78**, 1–8 (2008).
21. Aykol, M. & Wolverton, C. Local environment dependent GGA+U method for accurate thermochemistry of transition metal compounds. *Phys. Rev. B.* **90**, 1–18 (2014).
22. Friedrich, R. et al. Coordination corrected ab initio formation enthalpies. *npj Comput. Mater.* **5**, 1–32 (2019).
23. Isaacs, E. B. & Wolverton, C. Performance of the strongly constrained and appropriately normed density functional for solid-state materials. *Phys. Rev. Mater.* **2**, 1–11 (2018).
24. Hautier, G., Ong, S. P., Jain, A., Moore, C. J. & Ceder, G. Accuracy of density functional theory in predicting formation energies of ternary oxides from binary oxides and its implication on phase stability. *Phys. Rev. B* **85**, 155208 (2012).
25. Furness, J. W., Kaplan, A. D., Ning, J., Perdew, J. P. & Sun, J. Accurate and numerically efficient r^{2}SCAN meta-generalized gradient approximation. *J. Phys. Chem. Lett.* **11**, 8208–8215 (2020).
26. Perdew, J. P. et al. Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008).
27. Sun, J., Ruzsinszky, A. & Perdew, J. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **115**, 1–6 (2015).
28. Zhang, Y. et al. Efficient first-principles prediction of solid stability: towards chemical accuracy. *npj Comput. Mater* **4**, 9 (2018).
29. Yang, J. H., Kitchaev, D. A. & Ceder, G. Rationalizing accurate structure prediction in the meta-GGA SCAN functional. *Phys. Rev. B* **100**, 035132 (2019).
30. Sun, W. et al. Thermodynamic routes to Novel Metastable Nitrogen-Rich Nitrides. *Chem. Mater.* **29**, 6936–6946 (2017).
31. Park, J.-S. Comparison study of exchange-correlation functionals on prediction of ground states and structural properties. *Curr. Appl. Phys.* **22**, 61–64 (2021).
32. Bartel, C. J., Weimer, A. W., Lany, S., Musgrave, C. B. & Holder, A. M. The role of decomposition reactions in assessing first-principles predictions of solid stability. *npj Comput. Mater.* **5**, 4 (2019).
33. Mejia-Rodriguez, D. & Trickey, S. B. Deorbitalized meta-GGA exchange-correlation functionals in solids. *Phys. Rev. B* **98**, 115161 (2018).
34. Stevanovic, V., Lany, S., Zhang, X. & Zunger, A. Correcting density functional theory for accurate predictions of compound enthalpies of formation: Fitted elemental-phase reference energies. *Phys. Rev. B* **85**, 1–12 (2012).
35. Persson, K. A., Waldwick, B., Lazic, P. & Ceder, G. Prediction of solid-aqueous equilibria: Scheme to combine first-principles calculations of solids with experimental aqueous states. *Phys. Rev. B* **85**, 235438 (2012).
36. Ong, S. P. et al. Python Materials Genomics (pymatgen): a robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
37. Brumleve, T. R. Preparation, vapor pressure and thermochemistry of tin(II) bromide. *ECS Proc. Vol.* **1992-16**, 50–65 (1992).
38. Levin, I. Nist inorganic crystal structure database (icsd) (2020).
39. Patra, B., Jana, S., Constantin, L. A. & Samal, P. Correct structural phase stability of FeS_2, TiO_2, and MnO_2 from a semilocal density functional. *J. Phys. Chem. C.* **125**, 4284–4291 (2021).
40. Sun, W. et al. The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2**, e1600225 (2016).
41. Greenaway, A. L. et al. Ternary nitride materials: fundamentals and emerging device applications. *Annu. Rev. Mater. Res.* **51**, 591–618 (2021).
42. Sun, W. et al. A map of the inorganic ternary metal nitrides. *Nat. Mater.* **18**, 732–739 (2019).
43. Arca, E. et al. Zn_2SbN_3: Growth and characterization of a metastable photoactive semiconductor. *Mater. Horiz.* **6**, 1669–1674 (2019).
44. Greenaway, A. L. et al. Combinatorial synthesis of magnesium tin nitride semiconductors. *J. Am. Chem. Soc.* **142**, 8421–8430 (2021).
45. Caskey, C. M., Richards, R. M., Ginley, D. S. & Zakutayev, A. Thin film synthesis and properties of copper nitride, a metastable semiconductor. *Mater. Horiz.* **1**, 424–430 (2014).
46. Caskey, C. M. et al. Semiconducting properties of spinel tin nitride and other IV_3N_4 polymorphs. *J. Mater. Chem. C.* **3**, 1389–1396 (2015).
47. Mis, A., Lany, S., Brennecka, G. L. & Tamboli, A. Exploring the phase space of Zn_2SbN_3, a novel semiconducting nitride. *J. Mater. Chem. C.* **9**, 13904–13913 (2021).
48. Patel, A. M., Nørskov, J. K., Persson, K. A. & Montoya, J. H. Efficient pourbaix diagrams of many-element compounds. *Phys. Chem. Chem. Phys.* **21**, 25323–25327 (2019).
49. Wang, Z., Guo, X., Montoya, J. & Nørskov, J. K. Predicting aqueous stability of solid with computed pourbaix diagram using SCAN functional. *npj Comput. Mater.* **6**, 160 (2020).
50. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
51. 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).
52. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
53. Wisesa, P., McGill, K. A. & Mueller, T. Efficient generation of generalized Monkhorst-Pack grids through the use of informatics. *Phys. Rev. B* **93**, 1–10 (2016).
54. Ong, S. P. et al. 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**, 209–215 (2015).
55. Mathew, K. et al. Atomate: a high-level interface to generate, execute, and analyze computational materials science workflows. *Comput. Mater. Sci.* **139**, 140–152 (2017).

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Stephan Lany (National Renewable Energy Laboratory) for contributing structures in the Zn–Sb–N system, Julia Yang (University of California, Berkeley) for contributing the $SnBr_2$ structures, and Matthew McDermott (University of California, Berkeley) for helpful discussions regarding construction of phase diagrams with a shifted nitrogen chemical potential. This work was intellectually led by the Materials Project, which is funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, under Contract no. DE-AC02-05-CH11231: Materials Project program KC23MP. Additional support was also provided by the Data Infrastructure Building Blocks (DIBBS) Local Spectroscopy Data Infrastructure (LSDI) project funded by the National Science Foundation (NSF) under Award Number 1640899. A.S.R. acknowledges support via a Miller Research Fellowship from the Miller Institute for Basic Research in Science, University of California, Berkeley.

AUTHOR CONTRIBUTIONS

R.S.K.: conceptualization, software, methodology, data curation, formal analysis, visualization, validation, writing—original draft, writing—review & editing. A.S.R.: data curation, visualization, formal analysis, validation, writing—original draft, writing—review & editing. A.G.: formal analysis, visualization. J.M.M.: software, data curation, validation. S.P.O.: methodology, writing—review & editing. A.J.: methodology, writing—review & editing. S.D.: conceptualization, methodology, writing—review & editing, supervision. M.K.H.: conceptualization, methodology, writing—review & editing, supervision. K.P.: conceptualization, methodology, writing—review & editing, supervision, funding acquisition, project administration.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-022-00881-w.

Correspondence and requests for materials should be addressed to Kristin A. Persson.

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.