

Benchmarking VASP parameters for the PBE and r²SCAN workflows

Aaron Kaplan

Summary and recommendations:

- For statics, an ISMEAR = -5 calculation should always be performed
 - This often gives reasonable forces in metals, although it shouldn't in principle
 - If accurate forces in metals are desired, *a second* ISMEAR = 0 calculation should be performed
- We need to carefully ensure that the GGA tag is never specified when METAGGA is also specified
- LMAXMIX = 6 for all solids
 - This adds essentially no compute time (on average) over LMAXMIX = 2, 4
- LREAL = Auto is likely only safe near equilibrium
 - LREAL = False should be used preferentially
 - Both methods deliver comparable forces near equilibrium only, relaxations with LREAL = Auto are likely safe
 - *Speedups using LREAL = Auto are significant for equilibrated structures:* PBE and r²SCAN get a factor of 3-3.5 speedup with LREAL = Auto over False
 - *Speedups using LREAL = Auto are negligible for a wider set of static calculations for non-equilibrated solids*
- Parallelization: 1 CPU node on Perlmutter, KPAR = 2, NCORE = 16
 - Did not observe memory issues on this set, but KPAR = 1 could be used to lower memory usage
 - No multi-threading (OMP_NUM_THREADS=1)
 - No oversubscription

Remaining uncertainties:

- +U values need to be re-computed
 - Use original method of L. Wang, T. Maxisch, and G. Ceder [Phys. Rev. B **73**, 195107 (2006); DOI: 10.1103/PhysRevB.73.195107]?
 - Or use Guy's linear response method?
- vdW should be incorporated for materials with significant vdW
 - We found that r²SCAN+rVV10 *doesn't generally* improve r²SCAN's formation enthalpies, geometries, etc. [M. Kothakonda, A.D. Kaplan, *et al.*, ACS Mater. Au **3**, 102 (2023), DOI: 10.1021/acsmaterialsau.2c00059]
 - How to quantify "significant vdW"? When at least one nearest-neighbor distance is at least the sum of vdW radii?
- Version control of workflows
 - Should we label specific workflows for compatibility with "legacy" versions of MP?

ISM EAR

Motivation:

- ISMEAR controls the type of function used to broaden the occupation numbers near the Fermi level ϵ_F
 - The “true” occupation function at zero temperature is a step function
 - The high density of levels near ϵ_F for metals means a large density of k points are needed without smearing
 - The Methfessel-Paxton smearing method ($ISM EAR \geq 1$) expands the Dirac delta function in the basis of Hermite polynomials – it works well for metals, but can lead to unphysical occupation numbers (>1 or <0) – bad for gapped systems
 - Gaussian smearing works well generally provided the broadening $SIGMA \leq 0.05$ eV
 - The tetrahedron method with Blöchl corrections ($ISM EAR = -5$), which we usually use for single points, is not variational with respect to the partial occupancies in a metal
- In the context of the MatPES project, Shyue Ping (SP) expressed a desire to obtain accurate forces from static single points
- Alex Ganose wants to initialize all atomate2 relaxations with Gaussian smearing ($ISM EAR = 0$), but a *very large* smearing width of 0.2 eV
- SP wanted a benchmark of tetrahedron smearing compared with Gaussian smearing ($ISM EAR = 0$) with a small smearing width
- His students did a test for ~462 structures and found ~1 eV absolute deviations between the two methods
- This discrepancy can be extremely large on the scale of thermodynamic properties, 1-5 meV/atom

Method

- Having completed 1,011 formation enthalpy calculations for PBE and r²SCAN, we already have a baseline of $ISM EAR = -5$ calculations
- We took the relaxed structures from the second relaxation in each workflow as input to a static job with ($ISM EAR = 0$, $SIGMA = 0.05$ eV)
- All compounds and elemental structures were used

Results and Discussion

- After removing jobs which failed, or had incorrect smearing because of atomate2’s default behavior (specifying a previous calculation directory can override the user’s INCAR settings), have 1,070 PBE calcs and 1,090 r²SCAN calculations

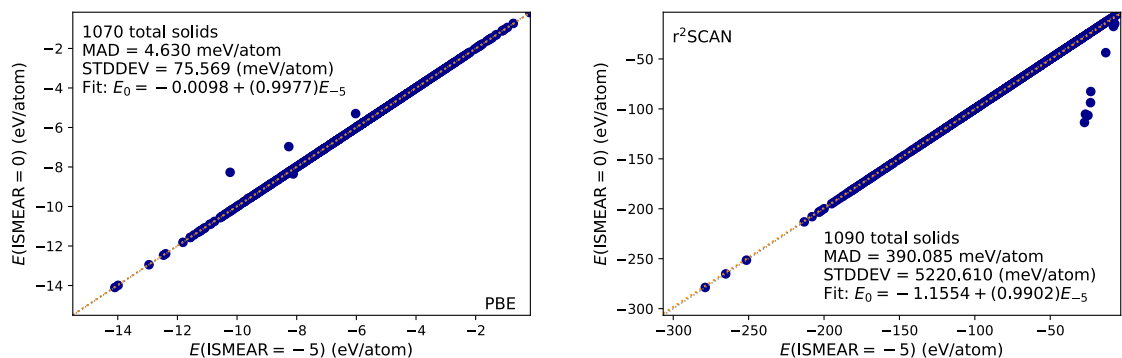
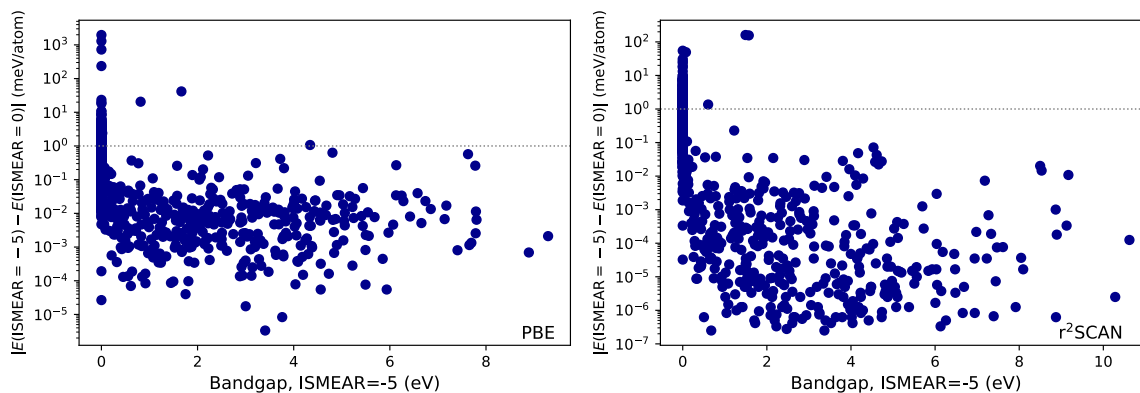


Figure 1. PBE (left) and r²SCAN (right) total energies (extrapolated to zero smearing width) for ISMEAR = -5, 0. The line of best fit is the dotted, orange line. The line of perfect agreement is the dotted gray line. The units of the intercept in the line of best fit are eV/atom, and the slope is dimensionless.

- Fig. 1 compares the energies obtained from both smearing methods
 - While the PBE energies broadly agree to about 5 meV/atom, this average discrepancy is significant for the scale of thermodynamic properties
 - More, the standard deviation for PBE is quite large
 - For r²SCAN, the deviations are even larger, with more significant outliers



- Fig. 2 plots the difference in total energies as a function of the bandgap obtained with ISMEAR = -5
 - Presume ISMEAR = -5 will give more accurate bandgap
 - As expected, differences are primarily for zero-gap metals, and mostly drop beneath 1 meV/atom threshold for gapped systems

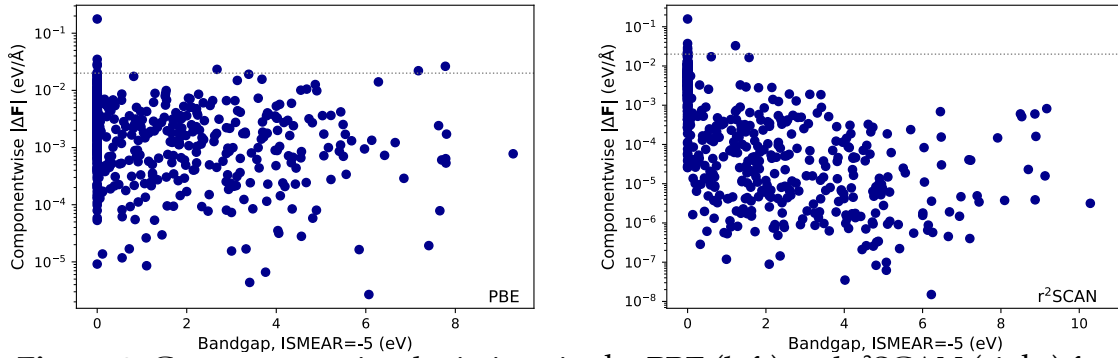


Figure 3. Component-wise deviations in the PBE (left) and r²SCAN (right) forces as a function of the ISMEAR = -5 (GGA = PS) bandgap. The dotted horizontal line is 0.02 eV/Å, which is used as the force criterion for EDIFFG.

- Fig. 3 plots the deviations in forces as a function of the bandgap
 - Let $\Delta\mathbf{F}(\mathbf{R}_k) = \mathbf{F}(\mathbf{R}_k; \text{ISMEAR} = -5) - \mathbf{F}(\mathbf{R}_k; \text{ISMEAR} = 0)$ be the difference in force evaluated using the two values of ISMEAR for the atom located at \mathbf{R}_k
 - Define the component-wise sum of forces on the M atomic sites as

$$|\Delta\mathbf{F}| = (3M)^{-1} \sum_{k=1}^M \sum_{j=1}^3 |\Delta F_j(\mathbf{R}_k)|$$
 - Curiously, both ISMEAR methods *for both* PBE and r²SCAN yield forces of comparable accuracy *even for metals*
 - Only a few outliers, mostly metals or small gap
 - Most forces lie below 0.02 eV/Å threshold we use for EDIFFG in the r²SCAN workflow

What happens when you specify both GGA and METAGGA tags?

- Because of how atomate2 works, when a previous directory to copy output files from (POSCAR, CHGCAR, WAVECAR, etc.), it also copies the INCAR settings
- Thus all the r²SCAN workflow calculations we did had a GGA = PS (PBEsol) tag propagated through to the meta-GGA relaxation and static from the initial PBEsol GGA relaxation
- A constant energy shift is introduced, which can change total energies *by an order of magnitude*
- Figs. 4 – 6 are analogs of the r²SCAN panels in Figs. 1 – 3, but the ISMEAR = 0 calculation *omits* the GGA tag
- *Forces may also be affected* as seen in Fig. 5
- A sanity check showed that setting GGA = PE (PBE) or "--" (use the functional from the pseudopotential) with PBE pseudopotentials gave no changes in energies or forces

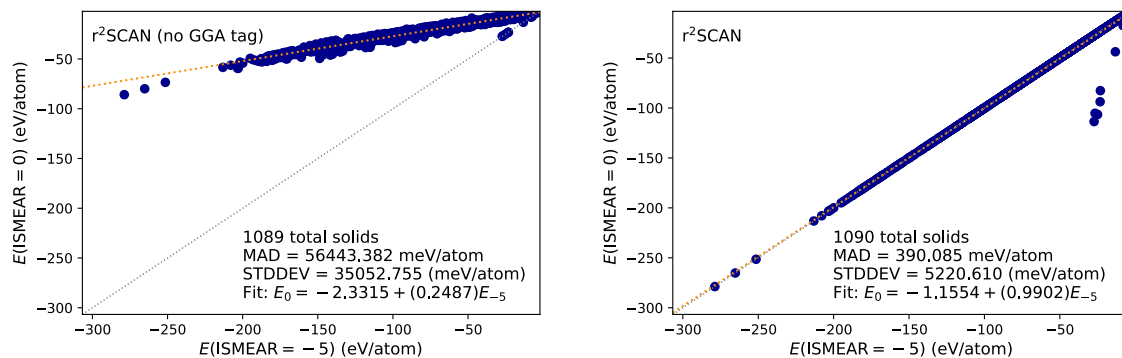


Figure 4. Total r²SCAN energies (extrapolated to zero smearing width) for ISMEAR = 0 and -5 (GGA = PS). The left panel uses no GGA tag (GGA = --), and the right uses GGA = PS, consistent with the ISMEAR = -5 calculation. The line of best fit is the dotted, orange line. The line of perfect agreement is the dotted gray line. The units of the intercept in the line of best fit are eV/atom, and the slope is dimensionless.

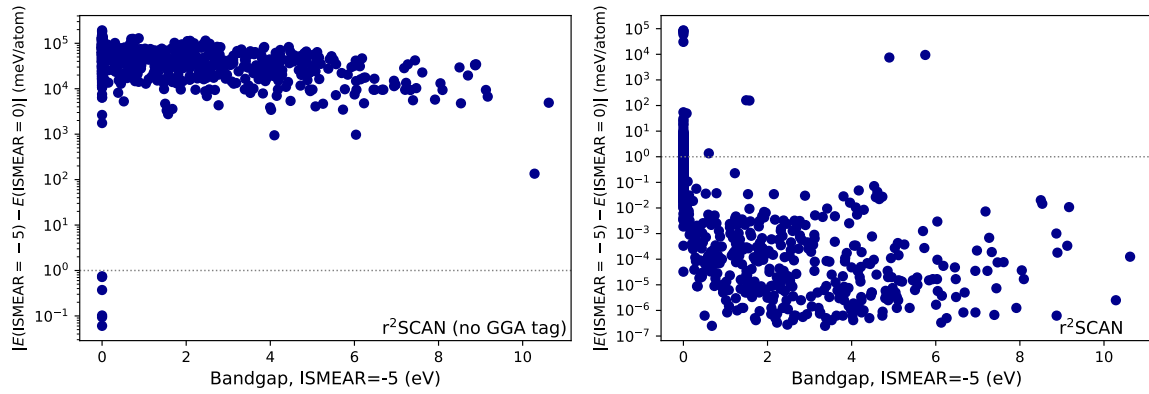


Figure 5. Difference in $r^2\text{SCAN}$ total energies obtained with $\text{ISMEAR} = -5$ (GGA = PS) and 0, plotted as a function of the bandgap (obtained with $\text{ISMEAR} = -5$). Left: no GGA tag (GGA = --); right: GGA = PS consistent with the $\text{ISMEAR} = -5$ calculation. The dotted horizontal line is 1 meV/atom.

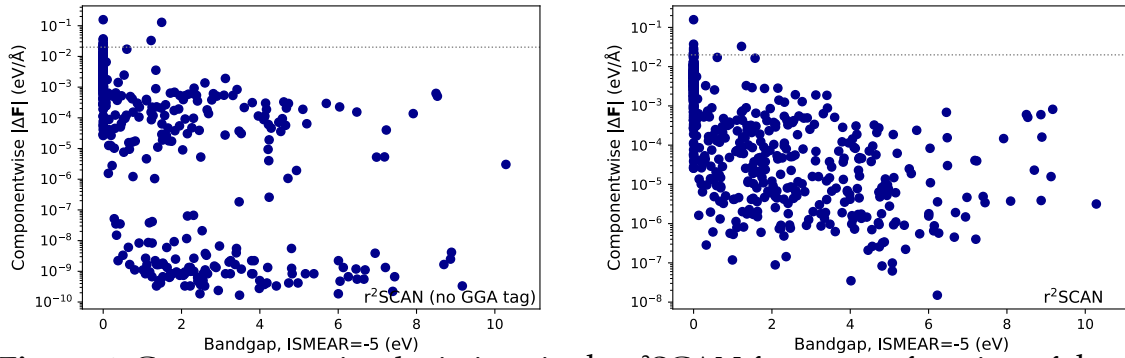


Figure 6. Component-wise deviations in the $r^2\text{SCAN}$ forces as a function of the $\text{ISMEAR} = -5$ (GGA = PS) bandgap. Left: $r^2\text{SCAN}$ with no GGA tag set; Right: $r^2\text{SCAN}$ with GGA = PS, consistent with the $\text{ISMEAR} = -5$ calculation. The dotted horizontal line is $0.02 \text{ eV}/\text{\AA}$, which is used as the force criterion for EDIFFG.

LMAXMIX

Motivation:

- LMAXMIX is used to expand the charge densities up to spherical harmonics of order $Y_{\ell m}$, with $\ell = \text{LMAXMIX}$
 - $\ell = 2$ for p -block, 4 for d -block, 6 for f -block, etc.
- The VASP manual suggests that LMAXMIX is *only* important for DFT + U
- Andrew Rosen previously found that the CHGCAR's produced by setting LMAXMIX too low for d -block compounds
 - Looked at calculations which had a non-self-consistent calculation (NSCF)
 - Found that NSCF calcs evaluated on a CHGCAR with too low LMAXMIX gave bandgaps, magmoms, etc. in poor agreement with SCF calc
- Matthew Kuner was curious if p -block elements with full d -shells require LMAXMIX = 4 as well
- Broaden scope to study:
 - How does compute time change with LMAXMIX?
 - Are there any cases where LMAXMIX = 8 are needed?
 - How does LMAXMIX affect NSCF energies, forces, stresses, gaps?
 - How do SCF energies converge with LMAXMIX?
 - Selected representative set of solids with open and closed d and f shells on the hull: Al_2O_3 , SnO , MnO_2 , FeO , Fe_2O_3 , UO_2 , CeS , PbO , and BiF_3 , and their elemental constituents

Results and discussion:

- There are essentially no differences in compute time when measured relative to the number of sites in the cell

| LMAXMIX | Avg. compute time (min./site) | Std. dev. (min./site) |
|---------|-------------------------------|-----------------------|
| 2 | 14.84 | 19.62 |
| 4 | 14.92 | 19.70 |
| 6 | 14.88 | 19.53 |
| 8 | 14.98 | 19.78 |

- Fig. 7 plots the difference in self-consistent and non-self-consistent (initialized from the same CHGCAR) total PBE energies
 - Only for LMAXMIX = 6 are *all* calculations converged to 1 meV/atom tolerance
 - Includes d -block elements
 - Same goes for convergence of self-consistent total energies wrt E(LMAXMIX=8)
- Significant outlier in all cases is mp-35 (elemental Mn) which requires LMAXMIX = 6 for well-converged charge density
- Fig. 8 plots the forces as a function of LMAXMIX, and Fig. 9 plots the bandgap. These are less sensitive to LMAXMIX
- Appears that all quantities are well-converged for LMAXMIX = 6, regardless of elemental block

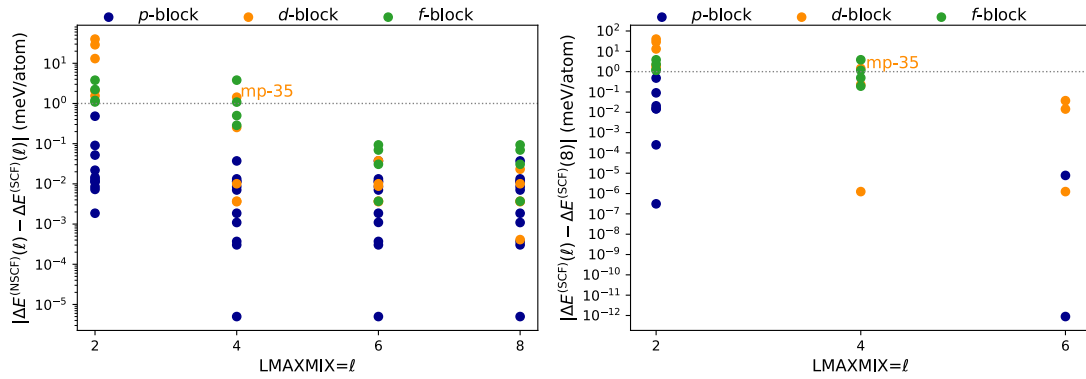


Figure 7. Left: difference in total energies evaluated non-self-consistently and self-consistently as a function of LMAXMIX. Right: convergence of self-consistent total energy wrt $E(LMAXMIX=8)$ as a function of LMAXMIX. The horizontal dotted line marks 1 meV/atom.

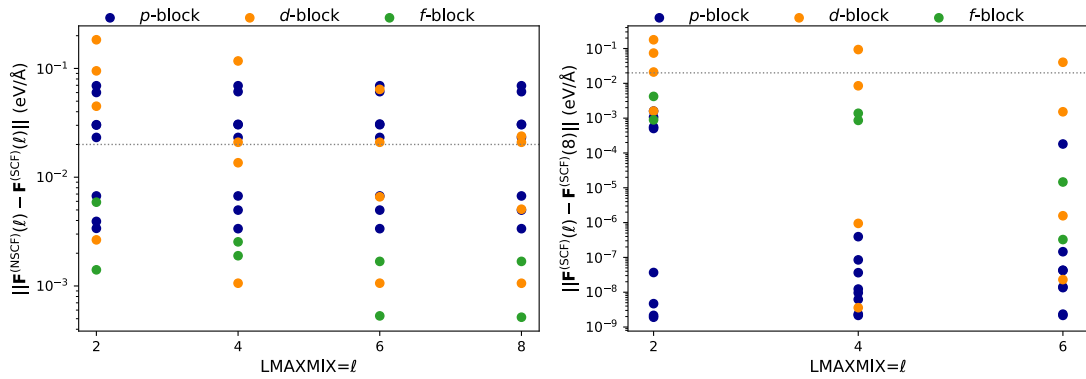


Figure 8. Convergence of RMS deviations in the forces evaluated between self-consistent and non-self-consistent calculations (left), and for self-consistent calculations wrt $E(LMAXMIX=6)$ (right). The horizontal dotted line marks 0.02 eV/Å.

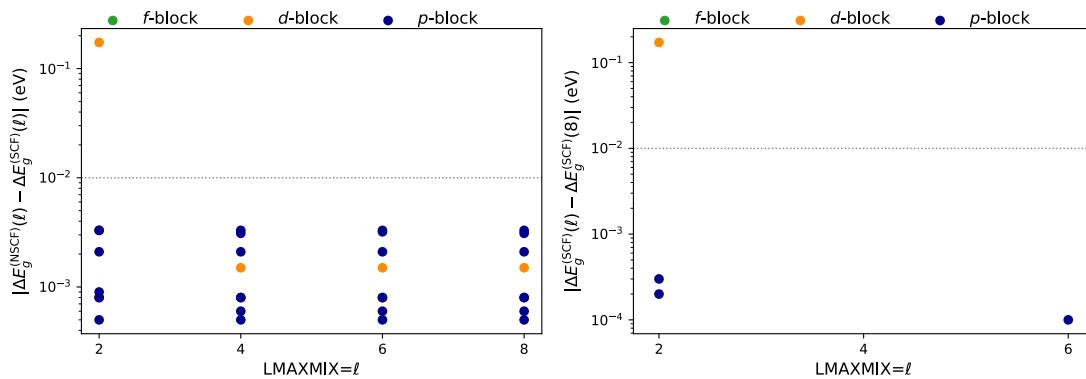


Figure 9. Same as Fig. 4, but for the bandgap.

LREAL

Motivation:

- The PAW PSPs require computation of projector operators, which can either be evaluated in real- or reciprocal-space
 - In real-space, the projectors can be affected by aliasing (high-frequency Fourier components are downfolded, yielding noise)
 - In reciprocal space, there is no aliasing issue, but the number of operators scales with the size of the system
- VASP recommends using an optimized real-space projection (LREAL = Auto) for systems with ≥ 30 atoms, and reciprocal-space projection (LREAL = False) otherwise
- Not clear that energies computed with LREAL = False (“more accurate”) and LREAL = Auto (“less accurate but cheaper”) are compatible

Results and discussion:

- Reporting results for 1,090 solids with PBE, and 1,089 solids with r²SCAN from the initial set
- Broadly speaking, LREAL = Auto and False yield essentially the same energies to within 1 meV/atom
- Fig. 10: some significant outliers hamper this trend, but they tend to be less frequent as system size grows beyond 20 atoms/cell – see Fig. 11
- From Fig. 11, it appears that setting LREAL = Auto is safe for cells with more than 30 atoms
- Fig. 12 demonstrates that the forces are much less sensitive to the choice of LREAL than the total energies, thus relaxations with LREAL = Auto are likely safe
- Timing information for the statics used in this comparison – see also Fig. 13

| LREAL = | False (min./site) | Auto (min./site) |
|---------------------|-------------------|------------------|
| PBE | 0.35 ± 0.28 | 0.12 ± 0.03 |
| r ² SCAN | 0.55 ± 1.49 | 0.16 ± 0.12 |

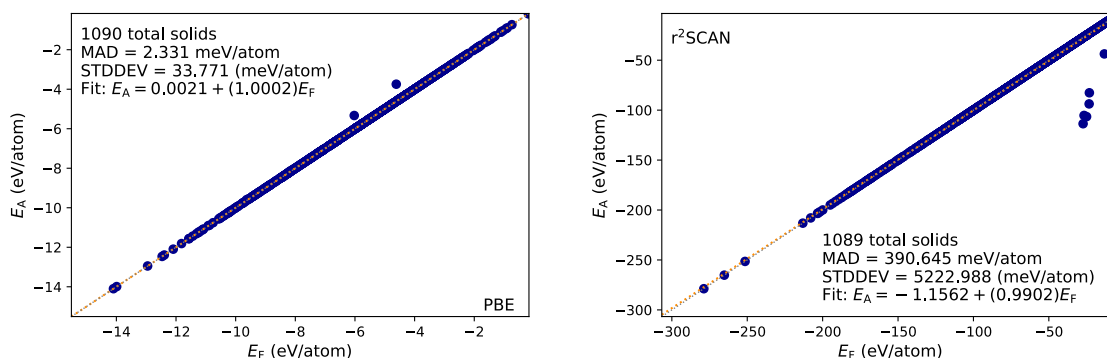


Figure 10. Total energies computed with LREAL = False E_F (horizontal) and with LREAL = Auto E_A (vertical), for PBE (left) and r²SCAN (right).

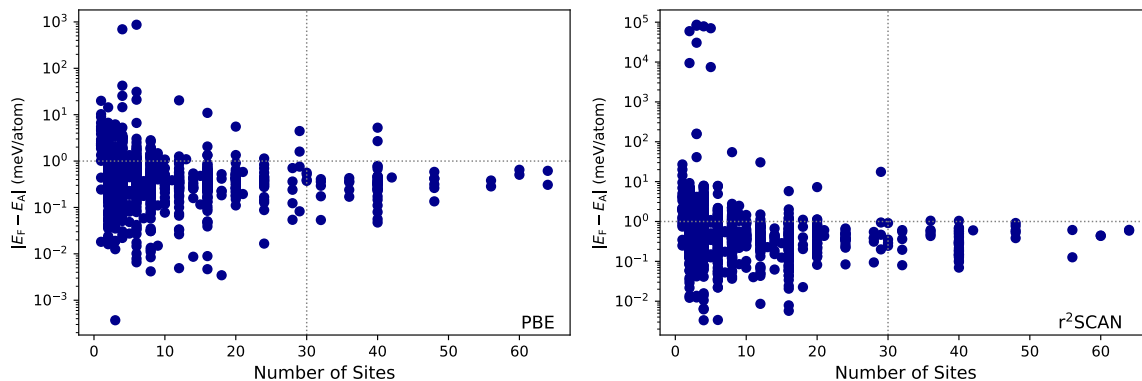


Figure 11. Difference in LREAL = False (E_F) and Auto (E_A) total energies (meV/atom) as a function of the number of atoms in the structure. Left: PBE, right: r^2 SCAN. The horizontal line marks 1 meV/atom, and the vertical line marks VASP's recommended cutoff for switching from LREAL = False (less than 30 atoms per cell) to Auto (at least 30 atoms per cell).

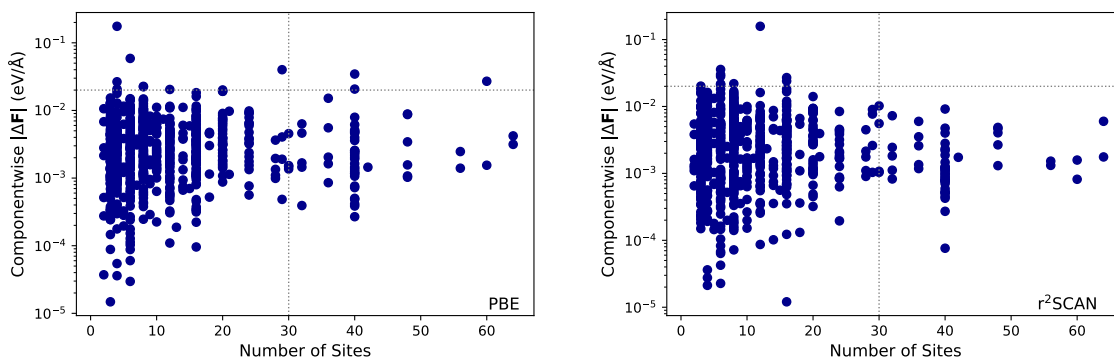


Figure 12. Component-wise deviation in the forces from LREAL = Auto and False. Left: PBE, right: r^2 SCAN. The forces are essentially unaffected by the choice of LREAL.

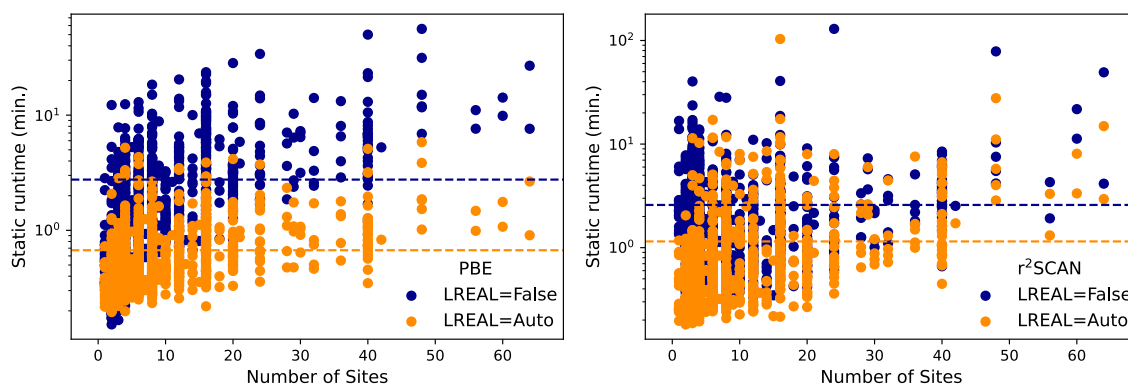


Figure 13. Total elapsed runtime for the PBE (left) and r^2 SCAN (right) statics using LREAL = False (dark blue) and Auto (orange). The average runtimes for each mode are indicated by horizontal dashed lines.

Non-equilibrated structures

Motivation:

- The previous tests involved structures that had been partly or fully relaxed by PBE
- While this gives useful information, it does not paint a complete picture of a typical WF for a completely novel material
- More, the previous set featured mostly systems with less than 30 atoms/cell
- Randomly selected structures from the ICSD which could not be matched to any entry in MP
 - Picked up to 5 structures with 1 atom/cell, 2 atoms/cell,...,100 atoms/cell → 480 candidate structures
 - Repeated the ISMEAR and LREAL validation tests with this set
 - Removed effects of other parameters in VASP: ENCUT = 680 eV, ENAUG = 1360 eV, LMAXMIX = 6, EDIFF = 10^{-6} eV for all calculations
 - Just “tight static” runs
 - For ISMEAR tests, LREAL = False
 - For LREAL test, ISMEAR = -5

Results and discussion:

- After removing failed runs, 392 ISMEAR and 363 LREAL comparisons could be made
 - Validated that anticipated settings were used during the run (i.e., runs where custodian changed ISMEAR were discarded)
- Conclusions are surprisingly different:
 - ISMEAR timing differences negligible: 0.14 ± 0.13 min/site for ISMEAR = -5 and 0.16 ± 0.16 min/site for ISMEAR=0
 - ***LREAL timing difference also negligible***: 0.13 ± 0.14 min/site for LREAL = False vs 0.16 ± 0.16 min/site for Auto
 - Energies consistent between ISMEARs and LREALs (Figs. 14 - 17)
- Fig. 18: Forces deviate strongly for metals and semiconductors between ISMEAR methods
 - Mostly insensitive to number of sites in a cell
- Fig. 19: Force deviations *grow* as a function of the number of sites in a cell for LREAL = Auto and False
 - Constant as a function of bandgap

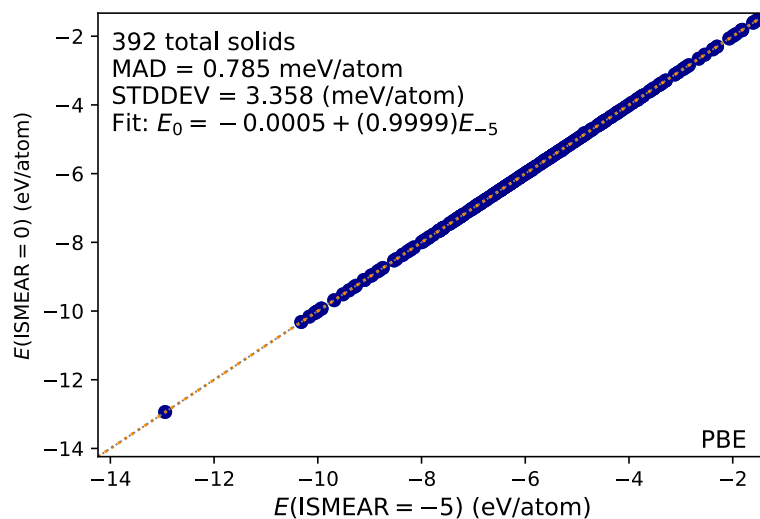


Figure 14. Comparison of total energies obtained for 392 ICSD structures which could not be matched to any entry in MP. Horizontal axis: ISMEAR = -5 (tetrahedron), vertical axis: ISMEAR = 0. The line of best fit is orange, the line of perfect agreement is dotted gray.

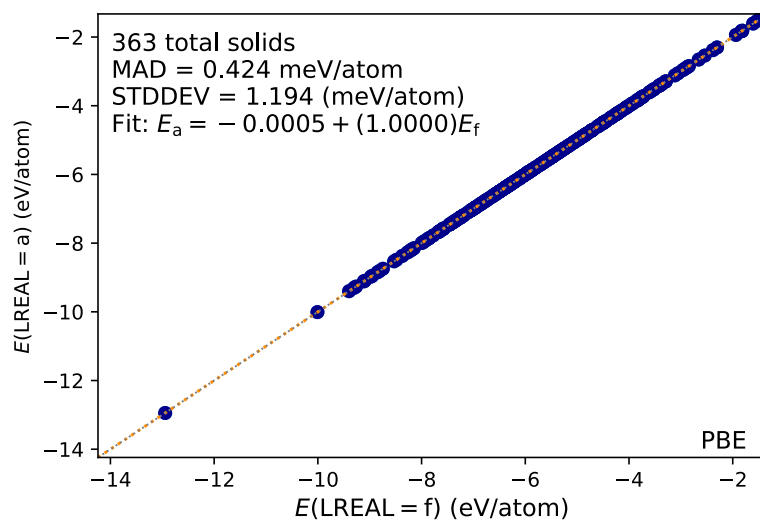


Figure 15. Same as Fig. 14, but for LREAL = False (horizontal) and Auto (vertical).

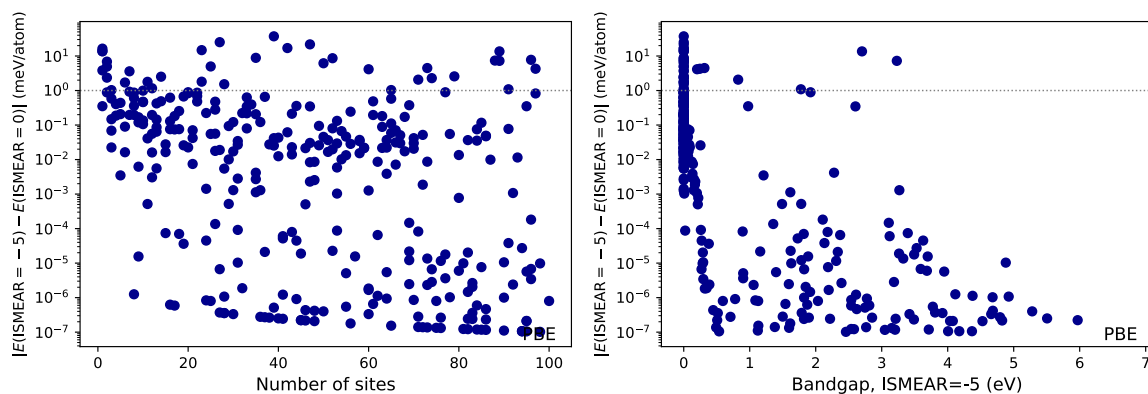


Figure 16. Plot of the total energy difference between ISMEAR = 0 and -5 as a function of the number of sites in a cell (left), or the bandgap (right) for non-equilibrated ICSD structures. The horizontal line indicates the 1 meV/atom acceptable uncertainty.

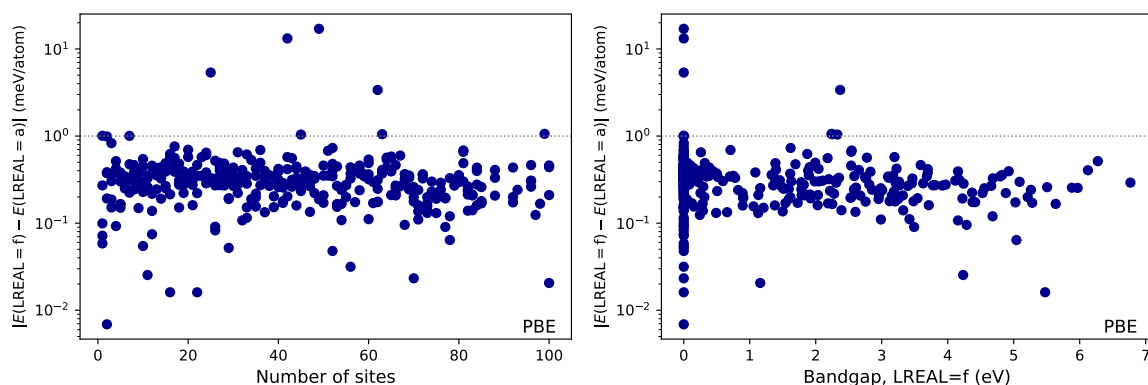


Figure 17. Same as Fig. 16, but comparing LREAL = False and Auto.

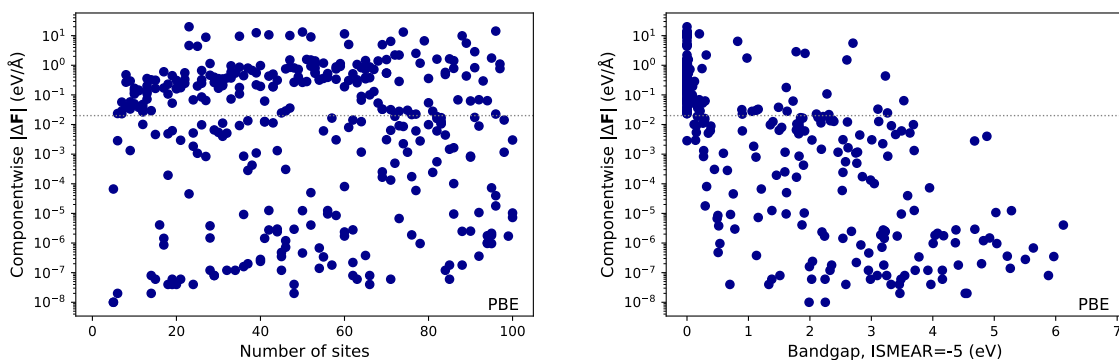


Figure 18. Componentwise deviation in the forces for non-equilibrated ICSD structures using ISMEAR = 0 and -5, as a function of the number of sites in a cell (left), or the bandgap (right). The horizontal dotted line indicates the 0.02 eV/Å threshold suggested for EDIFFG.

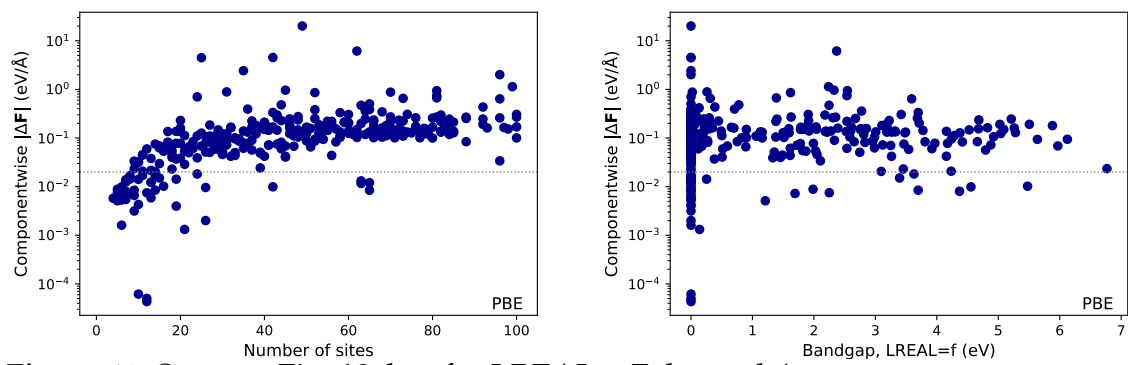


Figure 19. Same as Fig. 18, but for LREAL = False and Auto.