Active Learning Discovery of Oxidized and Active

IrOx Phases

Raul A. Flores,† Christopher Paolucci,‡ Ankit Jain,¶ Kirsten T. Winther,† Jose

Antonio Garrido Torres,† Muratahan Aykol,§ Jens K. Nørskov,¶ Michal Bajdich,∗,k and Thomas Bligaard∗,k

1. *SUNCAT Center for Interface Science and Catalysis, Department of Chemical*

*Engineering, Stanford University, Stanford 94305, California, USA*

1. *Department of Chemical Engineering, University of Virginia, Charlottesville, Virginia*

*22903, United States*

¶ *Department of Physics, Technical University of Denmark, Lyngby, Denmark*

1. *Toyota Research Institute, Los Altos, CA 94022, USA*

k *SUNCAT Center for Interface Science and Catalysis, SLAC National Accelerator*

*Laboratory, Menlo Park, CA 94025, USA*

E-mail: bajdich@slac.stanford.edu; bligaard@stanford.edu

# Abstract

Machine learning (ML) has revolutionized a number of scientific fields where it is possible to train models that are flexible enough to regress to data of interest while maintaining predictive power. Particular impact of ML surrogate models to have been in the field of materials science, where the bottle-neck imposed by the computational expense of Density Functional Theory (DFT) when applied to to vast compositionspace of bulk systems has been addressed. However, surrogate model applications to the structural space of bulk-crystals, i.e. polymorphs, are very scarce due to **XX**.

Herein, we report on an active learning ML methodology that searches for the most stable crystal structures of IrO2 and IrO3 by utilizing surrogate models which optimize within a candidate data set of crystal motifs sourced from publicly available materials databases. We demonstrate the efficacy of this AL-accelerated methodology by discovering 70 percent of the 10 most stable crystal structures for IrO3 and IrO3 with less than 50 DFT calculations. For IrO2, we show that while the bulk-rutile system is the globally stable polymorph, while more than **XX** unique structures is discovered within 0.2 eV per atom. For IrO3, we discover **XX** previously unknown polymorphs, e.g., *α*− AlF3 type and rutile-like IrO3 with stabilities lower than 0.2 per atom than anything known to date. With these results as inputs, we construct a new bulk Pourbaix diagram of the Ir-H2O system. We computationally test the proficiency of these phases towards oxygen-evolution reaction and find that the above stable IrO3 polymorphs have much higher activity that any IrO2. This work opens up an opportunity to materials/catalysts structural discovery on unprecedented scale.

# Introduction

Predicting the thermodynamically favorable crystal structures for an arbitrary inorganic system remains a challenging problem in computational material science. When simulations are used to guide the search for new materials, the stable and meta-stable crystal structures, i.e. polymorphs, above the convex hull of stability must be known in order to predict the material properties. Although there have been numerous examples in recent years of machine learning algorithms applied towards the prediction of formation energies of large ab-initio data sets, these data sets are biased towards common structures and varying composition space . For example, roughly half the entries (˜200,000) in The Open Quantum Materials Database (OQMD) correspond to ternary-alloy combinations in the same close-packed cubic structure. As such, these efforts have been primarily concerned with the enumeration of composition (elemental identity and stoichiometries) and less so with the exploration of structural diversity. The task of finding globally/locally stable crystal structures is equivalent to performing a global optimization (GO) routine within the highly dimensional potential energy surface (PES). Traditional approaches, such as simulated annealing, are only tractable for the most simple systems, such as metallic crystals which tend to adopt highly symmetric close-packed configurations, but is less suited for more complex materials. For instance, the class of structurally diverse metal-oxides, an important class of materials which tend to organize themselves into well-defined local coordination environments (octahedral, tetrahedral, etc.) which can assemble in a large variety of configurations with long-range order. In the past, various groups have put forth methodologies to address the structuraldiversity problem using computation for the MnO2 and VOx polymorph spaces, but most of these methodologies are limited by the fact that they are operating within the highly intractable/dimensional PES. Here we report on a crystal structure discovery algorithm that leverages machine learning surrogate models and an active learning framework to accelerate the discovery of novel crystal structures at fixed composition. The algorithm avoids operating in the highly dimensional 3*N*-space by leveraging nature’s propensity for symmetry by preparing data sets with a large degree of structural diversity at fixed composition.

Herein, we focus on the chemical space of iridium oxide polymorphs, an important class of materials with applications in electrochemistry. In particular, rutile-IrO2 (Ir[4+] oxidation state), is the most stable form of iridium-oxide at standard conditions, and is a well studied

electrocatalyst for the oxygen evolution reaction (OER).1–8 Previous studies on SrIrO3 electrocatalyst for the OER demonstrated that Sr leaching might leave behind a highly oxidized Ir (Ir[6+] for hypothetical IrO3) and it was argued as one possibly for observed high OER activity.1 Other groups also observed such dissociation of IrOx catalyst and subsequent formation of amorphous-like layer of unknown structure.9 Highly oxidized IrO3 phases as also formed as the terminal structure of LixIrO3 anodes.9 For above reasons, we focused our study to search for stable polymorphs in the standard IrO2 stoichiometry as well as higher oxidation state corresponding to IrO3, while neglecting the possibility of mixed IrO2(OH) phase. Purely octahedral IrO3 leads naturally to 100 percent corner sharing octahedra, where all terminal surface Ir-oxygens are potentially OER active sites. Furthermore, such pure corner sharing octahedral crystals are known from in other systems such fluorites and

chlorites.

In the first section, we define our prototype space and introduce the active-learning surrogate model. Next, we highlight the application of AL to the IrO2 and IrO3 prototype space. Here we discuss the acceleration/performance and practical limitations of this approach as well as the nature of the most stable polymorphs. Here, we also extract and analyze the rich structural information of our set. In the section 3, we construct a revised bulk Pourbaix diagram of the Ir-H2O system highlighting the importance of the IrO3 phases under OER. Finally, we construct thermodynamic OER volcano of most stable phases and discuss the

trends in activities.

# Results and discussion

## I. Candidate Space Generation and Active Learning Methodology

Here, we present a machine-learning based methodology for discovering new stable and metastable crystal structures. Our approach builds on the principles of surrogate active learning [Refs?], where a model is iteratively trained on available DFT data. The model predictions are used as a surrogate to the DFT energy evaluations, which are used to acquire new systems for DFT in the population based on an acquisition criteria. Active learning in conjunction with genetic algorithms has been demonstrated to successfully speed up materials discovery for alloy nanoparticles,10 and ...[Other refs?], structural optimizations,11 and transition-state calculations.12 The methodology consists of two steps (outlined in 1). The first step is the generation of the candidate space, which defines the inclusive list of all crystal structures to be screened through during the search routine. Since the initial candidates determines which structures that can ultimately be discovered, it is crucial to define a candidate space that is sufficiently diverse. The second part of the algorithm is the iterative active learning algorithm.

In Schema 1 we show the **overall process with the integrated active learning loop**. **We start with ...** . **We continue with with XXX...** The structures that comprise the candidate data sets for IrO2 and IrO3 were constructed by parsing for all bulk AB2 and AB3 structures in both the Materials Project13 and OQMD14 databases (TEMP AB2 and TEMP AB3). Structurally redundant systems were then removed via a space-group based structural classification scheme developed by Jain et al.15 The resulting data set is composed of 697 AB2 and 259 AB3 structural prototypes for which iridium and oxygen were replaced for the A and B sites, respectively. Finally, a coarse isotropic volume relaxation was performed to accommodate to account for the difference in atomic radii of Ir and O to the elements that originally comprised the structure in OQMD/MP. Additional details about our method can be found in the SI. Only candidates that were successfully optimized with DFT were ultimately included so that the model could be properly validated, which gives a final candidate data set of 448 and 258 AB2 and AB3 structurally unique polymorphs. The size of this dataset isn’t particularly large, but will serve to demonstrate the polymorph discovery routine as a proof of concept and is small enough that it is tractable to perform the DFT for all structures such that the performance of the model can be benchmarked.

The candidate data set was featurized using the Voronoi tessellation fingerprinting scheme developed by Ward et al.16 which produces a 271 feature vector for each material that are insensitive to isotropic expansions and contractions in a crystals lattice. Herein, we apply our active learning model to the IrO2 and IrO3 spaces separately, because we are interested in the most stable polymorphs at each stoichiometry. Constraining the candidate space to having uniform stoichiometry reduces the 271 feature vector to 101 non-zero variance features, thus reducing the dimensionality of our problem significantly. Further dimensionality reduction is achieved via a principle component analysis (PCA),17 which was used to reduce the remaining

101 features to 11. The active learning algorithm proceeds through iterative generations of ML training, prediction, and acquisition steps that are visualized in figure 3. To start, Gaussian process (GP) regression is used to train a regression model on a small seed set of DFT formation energies from randomly selected structures in the candidate space. The model is then used to predict the formation energies of the entire candidate space. Following this, the predictions are used to determine which systems to acquire, by minimizing the so-called GP-UCB acquisition function:

*U* = *µ* − *κσ,* (1)

where *µ* and *σ* is the predicted mean and standard deviation of the formation energy, and *κ* is a free parameter to tune the relative weighting between exploiting low formation energy systems (small *κ*) and exploring high uncertainty regions of the candidate space (large *κ*). In this work we attempt to trade off exploitation and exploration by weighting the predicted formation energy and the associated uncertainty to bias systems that are both low energy and high uncertainty. Here, *κ* is set to 1 which equally weights the energy and uncertainty. Once ranked, the N systems that minimize the acquisition function are selected for full DFT calculations, which are included in the training data of subsequent AL generations. The AL loop proceeds until convergence is achieved, which here is chosen to be the generation at which the structures within the range of metastability, here taken as 0.1 eV/atom, are unchanging over three consecutive generations.

## II. Application of Active Learning to the discovery of stable Ir-O polymorphs

We now turn our attention of the application of this active learning scheme to the discovery of most stable forms of IrO2 and IrO3. The algorithm is applied separately to both stoichiometries. Here, we report in detail the results for IrO3, the analogous results for IrO2 are shown in more detail in the SI.

Candidate Space Generation

Materials databases

Select desired stoicheometry

Filter for structural uniquness

Replace elements with desired

Pre-optimize cell volume

Final candidate structures

Seed with random DFT Active Learning Search

ML

Training

STOP

Predict ΔEf Yes No

New stable structures

found?

Acquisition/DFT

Figure 1: Process flow diagram for the active learning accelerated algorithm. The procedure is composed of (a) generation of the candidate set of considered crystal structures constructed from DFT materials databases and (b) iterative active learning surrogate search of the candidate space.

Figure 3a shows a sequence of plots at various generations of the active learning loop, starting with the initial generation with five randomly drawn candidates and ending with the 40th generation of the ALL (205 DFT computed structures out of the 248 total candidates). Each plot tracks the predicted (hollow grey) and DFT-derived (solid red) formation enthalpies, sorted from most to least stable. As the active learning loop acquires DFT data, the GP model becomes more accurate, as evidenced by the decreasing uncertainties when going from left to right. At the top of each subplot of 3 the identity of top ten most stable polymorphs is tracked, with the short grey line turning into a longer red line when the structure is acquired by the ALL. At the first generation the top ten structures are randomly distributed across the entire candidate space because the GP model hasn’t had enough training data to identify the most stable polymorphs as low energy systems. After only three generations (3a.ii) the GP model is sufficiently trained to have correctly identified all of the top ten systems as being low energy. Additionally, by the third generation (20 DFT calculations) 2/10 top systems have been acquired. After another 3 generations (3a.iii) (15 additional DFT calculations) the AL routine has successfully identified 7/10 of the top systems. Figure 3e plots the number of top ten structures acquired as a function of DFT calculations for the ALL with the GP-UCB acquisition function and a baseline random acquisition scheme. The results of figure 3e are averaged over independent runs of AL algorithm with the 1 sigma standard deviation between these runs shown. Overall, the GP-UCB runs outperform the random acquisition runs, with only 50 DFT calculations on average needed to discover 7 of the top 10 systems. This is compared to the 157 DFT calculations needed to compare 7/10 top systems for the random acquisition.

The low energy region of 3a.iii is shown in 3c. And the 6 most stable structural polymorphs are shown in 3d.

Having demonstrated the efficacy of the ALL materials discovery scheme we will now turn our attention to evaluating the performance of the GP regressive model of the final generation (trained on all available IrO3 DFT data). Figure 3b plots the GP model predicted formation enthalpy against the DFT-computed values for two special cases, 1.) predicting onto the features of the pre-optimized structures (grey), as is done in the regular operation of the ALL when acquiring new structures and 2.) predicting onto the post-DFT fingerprints (blue). It is evident from the parity plot in figure 3b that the Gaussian process model is doing an poor job of predicting the DFT formation energy of the candidate space using the preoptimized fingerprints, with an exceptionally poor MAE of 1.5 eV/atom. The model’s poor predictions are skewed towards the higher formation enthalpies, with the errors associated with low energy structures being much more robust. The same GP model does comparatively much better at predicting the formation energies of post-DFT optimized structures with an MAE of 0.2 eV/atom, which is expected since the post-DFT fingerprints directly corresponds to the DFT energies. This is to show that the GP model’s poor predictive capabilities is due to the large degree of structural reorganization that occurs after DFT relaxation of the pre-optimized structure. Structures that are initialized in high energy configurations will therefore have high predicted formation enthalpies, and will then reconfigure into a lower nearby configuration, resulting in lower final energy and a large discrepancy between the predicted and final energies. It’s interesting to consider why the ALL appears to perform so well as outlined previously (discovering 7/10 of the most stable candidates after only 35 DFT calculations). The reason for this is that the pre-optimized structures that are similar enough to the most stable final equilibrium structures will not restructure considerably, meaning that their predicted formation energies will be close enough (and low enough) to be quickly picked up by the acquisition criteria.

## III. Crystal coordination analysis of discovered phases

**There should be a paragraph which discussed the structural drift and performance/acceleration in more detail** I would use the updated version of Figure 2c.

Next, we describe the structural variety that is present in our data set of 1000 IrO2 and IrO3 polymorphs. The coordination environment package ChemEnv, developed by Waroquiers et. al.18 and implemented in pymatgen,19was used to assign M-O crystal filed coordination types (e.g. octahedral, square pyramidal, cubic, etc.) to each of the 1K structures in our combined data set. **Describe in two paragraphs what is shown in Figure 3 and in what the main consequences of these results**

Additional esoteric coordination environments were identified manually, see SI. The resulting distribution is included in figure TEMP, which plots the electronic energy and volume, both normalized on a per atom basis

## III. Electrochemical OER Application

We next performed *ab-initio* thermodynamic simulations to elucidate the electrochemical operational stability of IrOx and the OER activity of the four stable polymorphs (rutile-

IrO2, *α*-IrO3, rutile-IrO3, and *β*-IrO3) computed above (Fg. XYZ).

Figure 2: (a) Progress of the active learning algorithm at five subsequent generations. The AL generation and number of DFT training data at each generation is shown for each subplot. The enthalpy of formation is plotted, ordered from most to least stable, against all IrO3 candidates. Grey markers indicate predicted formation enthalpies from the ML model while red markers correspond to DFT-computed quantities. Error bars from the GP model corresponding to 1 sigma are shown for all predictions. The vertical lines at the top of each subplot are tracking the positions of the 10 most stable polymorphs at each generation and whether they have been acquired (red) or not (grey) by the AL routine. (b) Parity plot of the final ML models for IrO2 and IrO3 predicting on either the pre-optimized (grey) or the post-optimized structures of IrO2 and IrO3. (c) Zoomed inset of the 6th generation of the AL loop. (d) Crystal structures of 6 most stable IrO3 polymorphs. (e) The number of the top 10 most stable polymorphs of IrO3 that are discovered as a function of the number of DFT bulk relaxations, averaged over 5 independent runs of the AL algorithm using the GPUCB acquisition criteria (red) and a random acquisition method (grey). Error bars indicate the standard deviation over 5 runs. Red guide lines are displayed to show how many DFT calculations are needed to discover 7/10 of most stable polymorphs for the GP-UCB and random acquisition.

### Bulk Pourbaix

Figure **??** reports the IrOx Pourbaix diagram (E vs. pH) constructed with the following species: Ir, rutile-IrO2, *α*-IrO3, rutile-IrO3, *β*-IrO3, and an aqueous dissolved IrO[4−] species

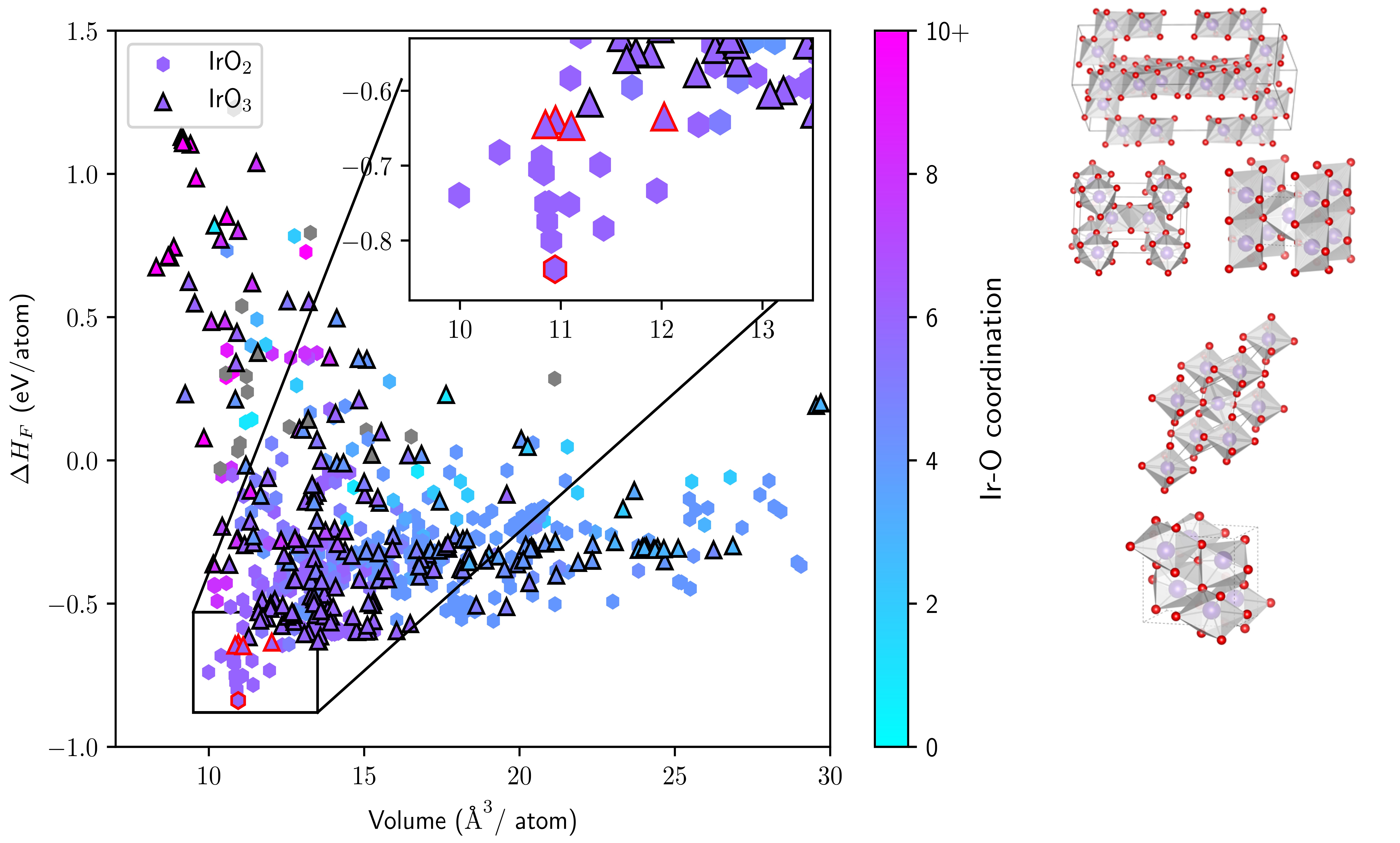


Figure 3: Enthalpy of formation for the 448 IrO2 (circles) and 258 IrO3 (crosses) structures in the candidate data set plotted against the volume per atom. Color overlays indicate the dominant coordination motifs as indicated by the legend. Select polymorphs systems are displayed around the plot area.

(See TEMP—SI for additional details). While and rutile-IrO2 are most stable at low bias, *α*-IrO3 is the most stable species under acidic conditions (pH *<*7) and in the bias region of interest for the OER ( 1.23 V vs. RHE). The stability regions of the metastable rutile-IrO3 and *β*-IrO3 phases are indicated by unfilled solid lines and appear (meta?)stable in the OER relevant region of the diagram. The similar formation energies (SI XYZ) for all three IrO3 species suggest some or all of these IrO3 phases may be present and are stable under OER

conditions.

### OER Surfaces and Activities

Fig. 5 summarizes the ~~major~~ results of the electrochemical activity and surface stability analysis, structure files and method details are reported in SI XYZ. Fig. 5 a.) reports the surface energy Pourbaix plots as a function of applied potential (at pH=0) for the four IrOx

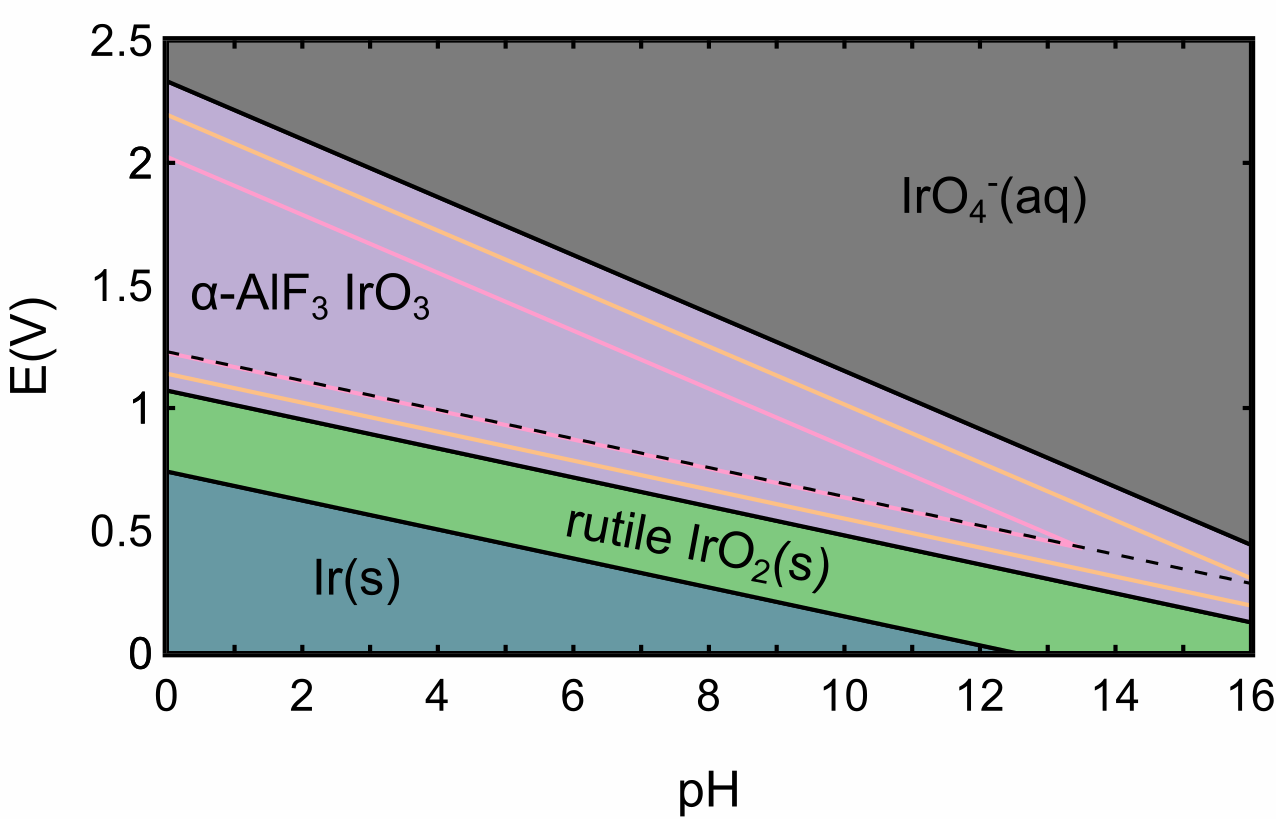


Figure 4: The revised Pourbaix diagram (electrochemical bulk phase stability) of the IrH2O system as a function of applied potential (vs. SHE) and pH. The most stable system studied (see Table XX in SI for a full list) are Ir-metal Ir(s) (blue), a rutile-IrO2 (green), and a dissolved IrO4[4−] (gray). Thee are compared to theIrO3 polymorphs, *α*-IrO3 (purple), rutile-IrO3 (orange), and *β*-IrO3 (pink). **I would like to add Srx IrO3 to show much less stable it is than alpha**. The water equilibrium line at U=1.23 V vs RHE shows the ideal onset of OER. This plot highlights the importance of the IrO3 for OER operational stability.

crystals of interest. The bulk phase limits of stability from figure TEMP are included at the bottom of each subplot. For each polymorph, XYZ specific facets were chosen from the highest intensity x-ray diffraction peaks from powder-diffraction spectra simulated in VESTA (TEMP insert vesta ref and ref SI xrd plot) and physical intuition. For each facet we computed the surface free energy for three coverages, bare, \*OH, and \*O. At modest overpotentials (eta 0.3 and potentials of 1.5 V vs RHE) the convex hull is populated solely by oxygen terminated surfaces. Consequently, we consider mainly oxygen terminated surfaces for the OER analysis.

The OER activity (expressed in terms of the limiting potential) for select oxygen terminated surfaces are shown in Fig. 5 as a function of the ∆*G*O − ∆*G*OH thermodynamic descriptor. The two rutile-IrO2 surfaces (100 and 110) bind the OER intermediates strongly, locating them at a theoretical limiting potential of xyz. The predicted overpotentials of our rutile-IrO2 systems are within the range of experimentally observed overpotentials found in literature. The three IrO3 polymorph surfaces all have a ∆*G*O − ∆*G*OH descriptor towards the top and right of the volcano, indicative of weaker binding energetics. This is evident from figure SI TEMP (scaling) which shows a clear distinction between the IrO2 and IrO3 polymorphs, with IrO3 binding on average TEMP eV weaker than IrO2. The best performing systems, including the (100), (110), and (211) facets of a-IrO3, b-IrO3 (101), and R-IrO3

(110), have overpotentials of 0.4 V vs RHE, a 0.2 V vs RHE improvement over the rutileIrO2 system. We note that the computed overpotentials for our rutile-IrO2 system differs from that reported in1 by 0.2 V. This discrepancy is due to our us of spin-polarization, which was neglected in Seitz et al., which strengthens the binding of IrO2.

# Conclusion

In conclusion, we have demonstrated an active-learning accelerated algorithm for the discovery of stable crystal polymorphs by searching through a candidate space of structurally distinct iridium-oxide phases. The algorithm can identify 7 of the 10 most stable polymorphs of IrO3 with only 35 DFT bulk relaxations and TEMP of the most 10 stable IrO2 polymorphs after TEMP DFT calculations. For IrO2, we find.... For IrO3 we find.... In the IrO2 space our search failed to uncover anything more stable than the rutile-IrO2 phase, while for IrO3 (a much less explored stoicheometry) we found several polymorphs phases that are predicted to be stable under OER conditions. We have analyzed the local and global structural coordination and revealed a large degree of structural diversity in our dataset (octahedral, tetrahedral, square-pyramidal, cubic, and square-planar) Although octahedral coordinations are energetically preferred TEMP TEMP. The most stable systems were used to construct a revised Pourbaix diagram of Ir-H2O system. Very importantly, we predict that IrO3 is the thermodynamically preferred phase under OER conditions. Finally, using thermodynamical approach to OER, we show, that surfaces of selected IrO3 have much higher relative activity that IrO2 due to presence of high valency Ir[6+] states. The 100 percent cornersharing octahedral structures feature maximum coverage of oxygens with optimal 2*p*-energy.

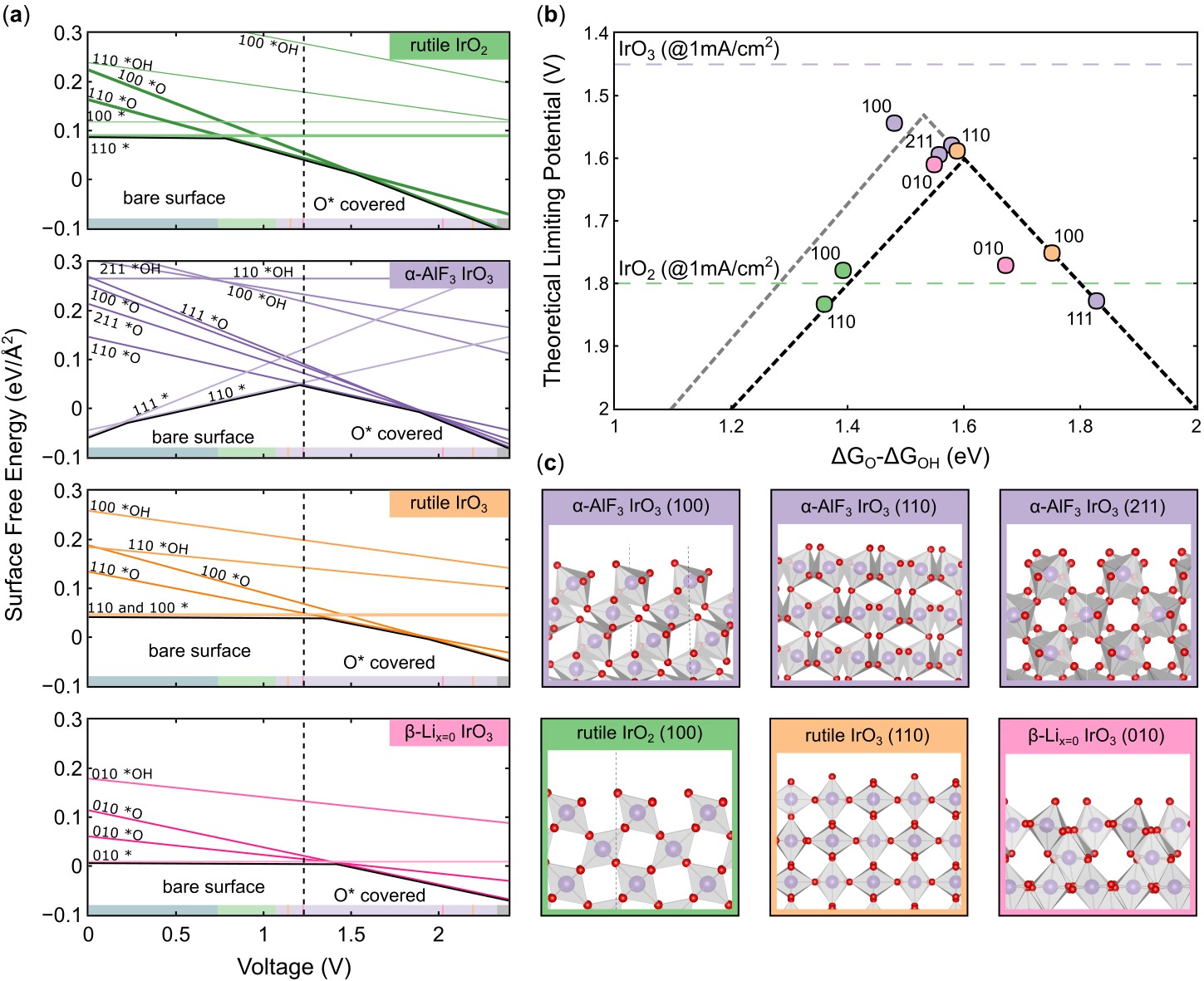


Figure 5: Summary of OER results for the four bulk structures of IrOx considered: rutileIrO2 (green), *α*-IrO3 (purple), rutile-IrO3 (orange), and *β*-IrO3 (pink). (a) Surface energy Pourbaix diagrams for each structure, with the surface energy of various facets and coverages shown as a function of applied potential. The bulk Pourbaix diagram’s bounds of stability at pH 0 are superimposed at the bottom of each subplot. (b) OER activity volcano for IrOx systems considered utilizing the ∆GO-∆GOH thermodynamic descriptor. The purple dotted line corresponds to the experimental limiting potential at 10 mA cm2 for IrO3,1 while the green band corresponds to the range of experimentally observed overpotentials for pristine IrO2 catalysts as reported in literature. (c) Select surface facets for the four IrOx crystal systems considered.

The OER results has broader implications for related OER systems. Overall, the AL ML has tremendous potential for discovery of structurally diverse systems particularly where the known diversity is currently very small (Highly oxidized oxides). Because our method provides readily available structural information which a necessary input for any characterization/simulation analysis TEMP We envision that ..... This opens up and new avenues of the materials/catalysis research with tailored structural properties.

Going forward we will improve

# Acknowledgement

This work was supported by the Toyota Research Institute.

JAGT and MB acknowledge the support by the U.S. Department of Energy, Office of Science, Office of Basic Energy Science, via Grant DE-SC0008685 to the SUNCAT Center of Interface Science and Catalysis.

The authors would like to acknowledge the use of the computer time allocation for the Transition metal-oxide and metal surfaces: applications and reactivity trends in catalysis at the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract

No. DE-AC02-05CH11231.

**Supporting Information Available**

# Active Learning ML Section

## Candidate space generation

The candidate for IrO2 and IrO3 were generated from existing experimental structures in the OQMD and Materials Project databases. There are TEMP unique AB2 structures (or multiples, e.g. A2B4) Of those we found 697 unique AB2 prototypes (unique SG/Wyckoff combination) in OQMD/MP

## Gaussian process regression model

Relevant details about the ML Gaussian process here

The Gaussian Process model utilized a rational quadratic kernel with variable length scales for each dimension of the feature space.

## Bulk polymorph DFT optimization

VASP PBE exchange correlation functional spin-polarized calculations plane-wave cutoff of 600 eV

A variable k-point mesh is used such that a k-point density of at least 20 k-points per reciprocal space dimension. All bulk systems were run through the following computational recipe to converge the equilibrium structure. The recipe has 3 distinct phases, and structures are only advanced to the next phase when the previous phase completes without error. 1. A ISIF 7 calculation to optimize only the volume (initial volume of cell may be really off) 2. 3 consecutive ISIF 3 relaxations to fully converge the lattice and atomic positions 3. A final ISIF 2, calculation to relax the atomic coordinates only to avoid errors associated with changing the cell volume with a fixed plane-wave cutoff basis The final ISIF 2 step is run with an electronic energy SCF convergence criteria of 1E-6 eV and the ionic relaxation has a tight force convergence criteria of 1E-3 eV/Angstrom

## Structural coordination motif identification

Several esoteric structural features were found in the DFT optimized structures and can be categorized as one of two types, legitimate structural motifs not characterized by the scheme of Waroquiers et. al.,18 and arguably non-physical structural artifacts, including: unassociated oxygen atoms - molecular oxygens in the unit cell

# Electrochemical OER Computational Methods

## Density Functional Theory Methods

All OER calculations were performed using density functional theory (DFT) implemented via the Vienna ab-initio simulation package (VASP) and utilizing the PBE exchange-correlation functional. Dipole corrections were imposed on all non-symmetric slabs. A 4x4x3 k-point mesh with gamma-point centered Monkshort-packing was used for all slabs. The plane-wave energy cutoff was 500 eV.

All slab calculations maintained a vacuum spacing of ¡15 A. All structures were relaxed utilizing a TEMP algorithm with a stop criteria being that all atoms satisfy a maximum force threshold of 0.02 eV/A.

## OER Thermodynamic Methodology

Procedure: - For the top/most stable bulk structures the following procedure was carried out

* Stable stoichiometric terminations were cut from the bulk Stable termination planeswere guesstimated via intuition, and the x-ray diffraction pattern tool from Vesta
* Electrochemical surface coverage was elucidated via a surface Pourbaix analysis Needto know the coverage of surface under operating conditions (¿1.23 V RHE)
* Thermodynamic/limiting potential analysis of the OER mechanistic pathway Volcanoplot, limiting potentials, etc.

## Surface Energy Pourbaix Methodology

Surface energy Pourbaix plots were constructed by calculating the surface energy of each slab by under standard conditions (V=0 and pH=0) and then utilizing the computational hydrogen electrode to compute the potential dependence of the surfaces.

Surface energy calculations were performed for various facets for slabs of increasing thickness. The bulk energy was then extracted by fitting the total energy of the slabs against the number of layers as explained in REF2. This was done to avoid common issues of surface energy divergence associated with using a separate bulk energy calculation.

The sensitivity of a given slab to an applied bias is dependent on the composition of the surface, in particular, the effect of coverage of electrolyte species which can deposit oxygen, hydrogen, and hydroxide species on the surface layers. These additional O and H atoms are not referenced to the atoms in the slab, but are instead referenced to the computational hydrogen electrode and water-splitting reaction. The equation for is as follows:

## OER Scaling Relations

Figure S1 shows the scaling relations between the adsorption free energies of the OER intermediate species for the IrOx systems studied herein. It can be seen clearly that the data points corresponding to the three IrO3 polymorphs are roughly 1 eV weaker binding than the rutile-IrO2 points. This generally weaker binding of the IrO3 stoichiometry is responsible for the observed improvement in theoretical activity. The ∆GOOH vs.∆GOH relationship is very close to the traditional “universal scaling relations”, demonstrating that our materials do not break the infamous ∆GOOH vs. ∆GOH scaling.

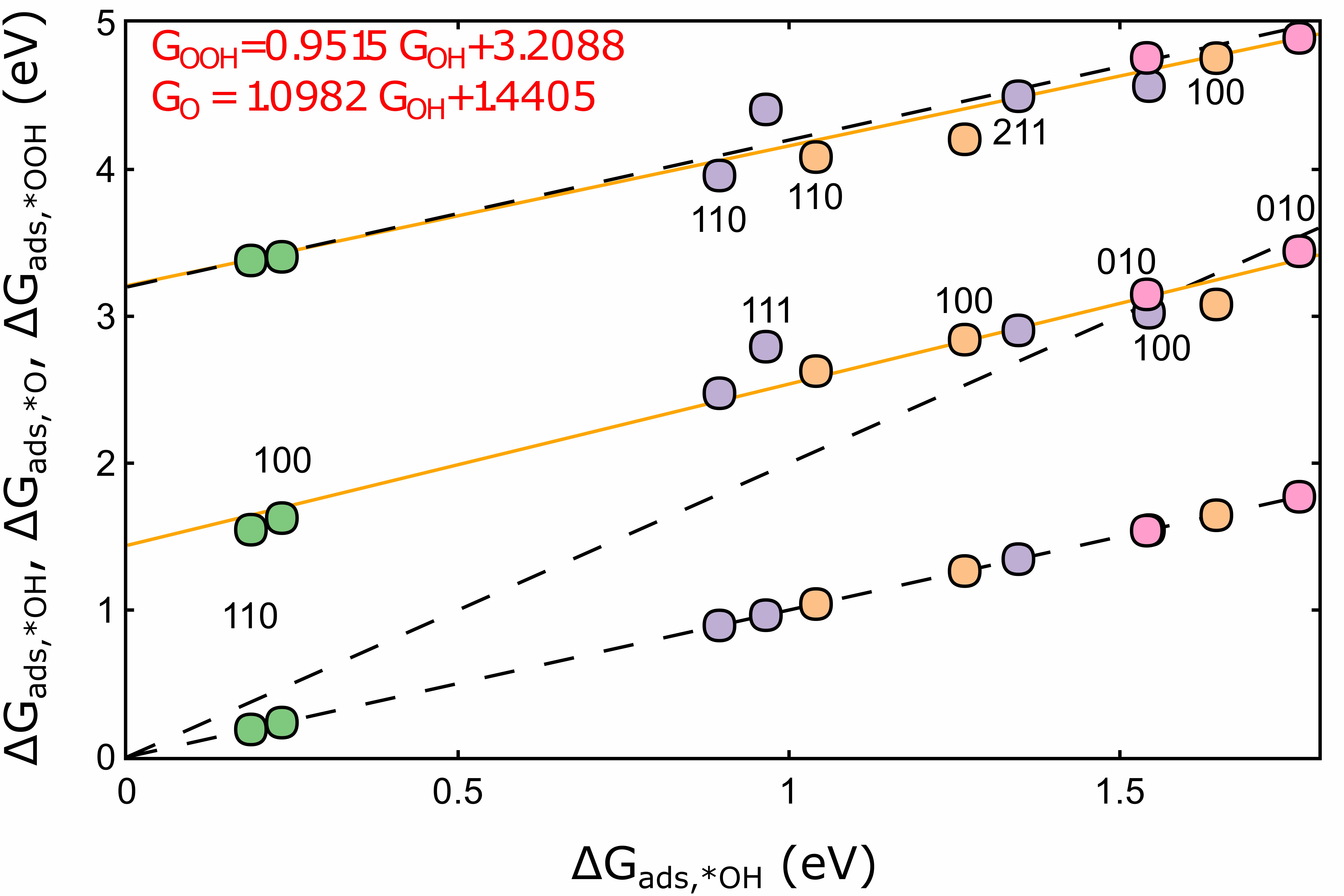


Figure S1: Relationship between the adsorption free energies of the three key OER intermediates (\*OH, \*O, \*OOH), with ∆GOH chosen as the dependent variable. Best fit lines are provided for ∆GOOH vs. ∆GOH and ∆GO vs. ∆GOH. Additionally, “universal scaling relations” for ∆GOOH vs. ∆GOH and ∆GO vs. ∆GOH are shown (black dotted lines) to emphasize our deviation from the traditionally reported scaling fits. The ∆GOH line is shown as guide to eye.

**Bulk Systems**

**Table of OER energetics**

# References

1. Seitz, L. C.; Dickens, C. F.; Nishio, K.; Hikita, Y.; Montoya, J.; Doyle, A.; Kirk, C.;

Vojvodic, A.; Hwang, H. Y.; Norskov, J. K.; Jaramillo, T. F. A highly active and stable IrOx/SrIrO3 catalyst for the oxygen evolution reaction. *Science* **2016**, *353*, 1011–1014.

1. Lee, Y.; Suntivich, J.; May, K. J.; Perry, E. E.; Shao-Horn, Y. Synthesis and activities of rutile IrO 2 and RuO 2 nanoparticles for oxygen evolution in acid and alkaline solutions. *Journal of Physical Chemistry Letters* **2012**, *3*, 399–404.
2. McCrory, C. C. L.; Jung, S.; Ferrer, I. M.; Chatman, S. M.; Peters, J. C.;

Jaramillo, T. F. Benchmarking Hydrogen Evolving Reaction and Oxygen Evolving Reaction Electrocatalysts for Solar Water Splitting Devices. *Journal of the American Chemical Society* **2015**, *137*, 4347–4357.

1. Trotochaud, L.; Ranney, J. K.; Williams, K. N.; Boettcher, S. W. Solution-cast metal oxide thin film electrocatalysts for oxygen evolution. *Journal of the American Chemical Society* **2012**, *134*, 17253–17261.
2. Danilovic, N.; Subbaraman, R.; Chang, K. C.; Chang, S. H.; Kang, Y. J.; Snyder, J.; Paulikas, A. P.; Strmcnik, D.; Kim, Y. T.; Myers, D.; Stamenkovic, V. R.; Markovic, N. M. Activity-stability trends for the oxygen evolution reaction on monometallic oxides in acidic environments. *Journal of Physical Chemistry Letters* **2014**, *5*, 2474–2478.
3. Carmo, M.; Fritz, D. L.; Mergel, J.; Stolten, D. A comprehensive review on PEM water electrolysis. *International Journal of Hydrogen Energy* **2013**, *38*, 4901–4934.
4. Miles, M.; Klaus, E.; Gunn, B.; Locker, J.; Serafin, W.; Srinivasan, S. The oxygen evolution reaction on platinum, iridium, ruthenium and their alloys at 80C in acid solutions. *Electrochimica Acta* **1978**, *23*, 521–526.
5. Beni, G.; Schiavone, L. M.; Shay, J. L.; Dautremont-Smith, W. C.; Schneider, B. S. Electrocatalytic oxygen evolution on reactively sputtered electrochromic iridium oxide films [7]. **1979**, *282*, 281–283.
6. Pearce, P. E.; Perez, A. J.; Rousse, G.; Sauban`ere, M.; Batuk, D.; Foix, D.; McCalla, E.; Abakumov, A. M.; Van Tendeloo, G.; Doublet, M. L.; Tarascon, J. M. Evidence for anionic redox activity in a tridimensional-ordered Li-rich positive electrode *β*-Li2IrO3. *Nature Materials* **2017**, *16*, 580–586.
7. Jennings, P. C.; Lysgaard, S.; Hummelshøj, J. S.; Vegge, T.; Bligaard, T. Genetic algorithms for computational materials discovery accelerated by machine learning. *npj*

*Computational Materials* **2019**, *5*, 46.

1. Hansen, M. H.; Torres, J. A. G.; Jennings, P. C.; Wang, Z.; Boes, J. R.; Mamun, O. G.; Bligaard, T. An Atomistic Machine Learning Package for Surface Science and Catalysis. *arXiv preprint arXiv:1904.00904* **2019**,
2. Torres, J. A. G.; Jennings, P. C.; Hansen, M. H.; Boes, J. R.; Bligaard, T. Low-Scaling

Algorithm for Nudged Elastic Band Calculations Using a Surrogate Machine Learning Model. *Physical review letters* **2019**, *122*, 156001.

1. Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The materials project:

A materials genome approach to accelerating materials innovation. *APL Materials* **2013**, *1*.

1. Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Ru¨hl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): Assessing the accuracy of DFT formation energies. *npj Computational Materials* **2015**, *1*.
2. Jain, A. PHYSICAL REVIEW B 98 , 214112 ( 2018 ) Atomic-position independent descriptor for machine learning of material properties. **2018**, *214112*, 1–7.
3. Ward, L.; Liu, R.; Krishna, A.; Hegde, V. I.; Agrawal, A.; Choudhary, A.; Wolverton, C. Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations. *Physical Review B* **2017**,
4. Tipping, M. E.; Bishop, C. M. Probabilistic principal component analysis. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* **1999**,
5. Waroquiers, D.; Gonze, X.; Rignanese, G. M.; Welker-Nieuwoudt, C.; Rosowski, F.; Go¨bel, M.; Schenk, S.; Degelmann, P.; Andr´e, R.; Glaum, R.; Hautier, G. Statistical analysis of coordination environments in Oxides. *Chemistry of Materials* **2017**, *29*,

8346–8360.

1. Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.;

Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials*

*Science* **2013**, *68*, 314–319.

# Graphical TOC Entry

