Supporting Information for "Maximal Predictability Approach for Identifying the Right Descriptors for Electrocatalytic Reactions"

Dilip Krishnamurthy,[†,¶] Vaidish Sumaria,[‡,¶] and Venkatasubramanian Viswanathan[*,†,‡]

†*Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

‡*Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

¶*Equal Contribution*

E-mail: venkvis@cmu.edu

1 Methods

Calculations were performed using the projector augmented-wave (PAW) method as implemented in the GPAW program package using the recently developed Bayesian Error Estimation Functional with van der Waals correlation (BEEF-vdW), which has built-in error estimation capability[1]. The exchange correlation uses an ensemble of exchange correlation functionals resulting in an ensemble of energies from which the uncertainty in the adsorption energies can be calculated. For the hydrogen evolution reaction, metal catalysts of 2 x 2 surface cell with 4 layers separated by 10 $\mathring{A}$ of vacuum and periodic in x-y direction were considered. The hydrogen intermediate was adsorbed on an fcc(111) site with a coverage of 1/4 monolayer. A 10 $\times$ 10 $\times$ 1 k-point grid was used for the calculations. Rutile oxide catalysts were used for both the oxygen evolution reaction and the chlorine evolution reaction. For rutile oxides, we consider a 2×1 surface unit cell and a 4×4×1 k-point grid. The surface of the unit cell contains two bridge and two cus sites. Adsorbates bind strongly on the bridge sites than on the cus sites and therefore the bridge site is always occupied with oxygen and inactive. All the OER and ClER intermediates were therefore adsorbed on the cus site. We consider a 1/2 monolayer (with respect to only the active cus sites) of the intermediates on the surface for both the reactions. Metal catalysts are used for the oxygen reduction reaction. Intermediates OH^* and OOH^* are modeled by including an explicit layer of water to account for hydrogen bonding on a 4-layered $\sqrt{3} \times \sqrt{3}$ configuration for metals and $2\sqrt{3} \times 2\sqrt{3}$ configuration for $Pt_3Ni(111)$ with 1/3 monolayer (ML) coverage. O^* is modeled on a 4 layered 2×2 configuration for metals and 2×3 configuration for $Pt_3Ni(111)$ in an fcc site with a 1/4 monolayer (ML) coverage. A $6 \times 6 \times 1$ k-point grid was used for the $2 \times 2 \times 4$ unit cell and the k-points are scaled according to the different unit cells used. For all the calculations, the bottom two layers were kept fixed and the top two layers with the adsorbates were allowed to relax with a force criterion of < 0.05 eV / $\mathring{A}$. Dipole correction was implemented in all calculations with metal catalysts. Spin-polarized calculations were carried out wherever necessary.

2 Adsorption Energy Distribution

Using BEEF-vdW functional, ensemble of adsorption energies for various intermediates involved is generated. We use the following methodology for estimating the combined overall error in the adsorption energies for various intermediates[2] (The methodology is explained using the adsorption energy of H* for hydrogen evolution as an example).

- First the ensemble of H* adsorption energies for a given metal "X" with respect to a reference system (one that minimizes the overall prediction error) is calculated. In the case of Hydrogen evolution, the reference chosen is Rh(111). This is given as:

  $$\mathrm{E_H(X|Rh(111)) = E_H(X) - E_H(Rh(111))}$$

- We then center the distribution around the mean, which is given as:

  $$\overline{\mathrm{E_H(X|Rh(111))}} = E_H(X|Rh(111)) - \langle E_H(X|Rh(111))\rangle$$

- This is carried out for all the catalysts ("X") considered and the combined distribution is constructed. The standard deviation of this combined distribution is the overall error in the adsorption energies (σ_H)

3 Hydrogen Evolution Reaction

3.1 Reaction Mechanism

$H^+ + e^- + * \rightarrow H^*$
$H^+ + e^- + H^* \rightarrow H_2(g) + *$

3.2 Calculation Details

The calculations were done on a 2×2 surface cell with 4 layers separated by 10 $\AA$ of vacuum. the slab is periodic in x and y direction. The hydrogen is adsorbed on a fcc(111) site. We considered surface coverage of 1/4 and 1 monolayer. The bottom two layers were fixed and the top two layers with the adsorbates are allowed to relax. All the structures were converged with a force criterion <0.05 eV/$\AA$. A $10 \times 10 \times 1$ k-point grid was used for the calculations. The adsorption free energy of hydrogen is given as:

$$\Delta E_H = \frac{1}{n}(E(surf + nH) - E(surf) - \frac{n}{2}E(H_2(g))$$

$$\Delta G_{H^*} = \Delta E_H + \Delta E_{ZPE} - T\Delta S_H = \Delta E_H + 0.24 \text{ (eV)}$$

where n is the number of H atoms used in the calculation; n = 1 represents a coverage of 1/4 and n = 4 represents a coverage of 1. The limit where proton transfer is exothermic ($\Delta G_{H^*} < 0$), the rate constant is independent of ΔG_{H^*} and the surface coverage is high. The limit where proton transfer is endothermic ($\Delta G_{H^*} > 0$), the reaction is activated by at least ΔG_{H^*} and proton transfer becomes difficult because hydrogen is unstable on the surface. The exchange current i_0 can be expressed in terms of the free energy of adsorption of hydrogen and charge transfer coefficient α using the following expression[3]:

$$i_0 = -ek_0(1 + \exp(|\Delta G_{H^*}|/(\alpha kT)))^{-1}$$

The uncertainty in the descriptor ΔG_{H^*} is found using the standard deviation of the combined ensemble distribution of the free energies found using the BEEF-vdW exchange correlation.

Figure S1 shows the combined adsorption energy distribution for H* calculated on Au, Ag, Pd, Pt, Rh, Ir, Ni, W, Co, Cu, Mo, Re, and Nb for fcc(111) facet. Its important to understand that the reference Rhodium is systematically chosen such that the uncertainty in the descriptor (σ_H) is minimized.

3.3 Uncertainty Propagation

Following methodology is used to propagate the uncertainty in the descriptor to the exchange current:

- Using the uncertainty in the descriptor (σ_H) calculated using the combined re-centered distribution of adsorption energies, we now approximate a given computed ΔG_{H^*} as

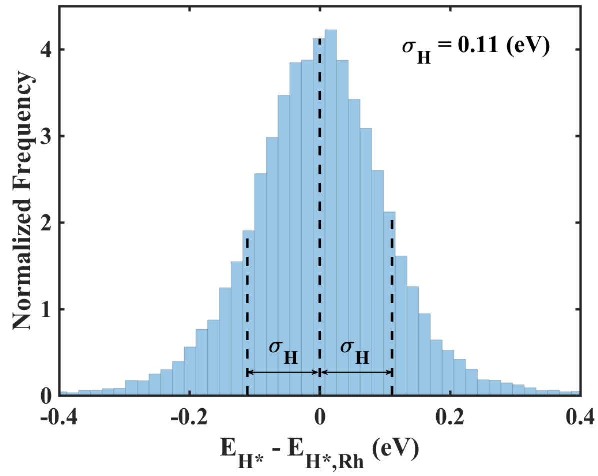

Figure S1: Plot of normalized frequency as a function of the adsorption energy of the intermediate H* relative to Rhodium. The standard deviation of this combined ensemble is σ_{H} = 0.11 eV.

a normal distribution with its mean, $\mu = \Delta\mathrm{G}_{\mathrm{H}^*}$ and standard deviation of σ_{H}. This distribution can be expressed as:

$$X \sim \mathcal{N}(\mu, \sigma_H^2)$$

- Using this distribution the probability distribution of the descriptor, $\Delta\mathrm{G}_{\mathrm{H}^*}$, can be given by the Gaussian distribution:

$$\mathrm{p_x}(\mathrm{x}|\mu, \sigma_{\mathrm{H}}^2) = \frac{1}{\sqrt{(2\pi\sigma_{\mathrm{H}}^2)}} \exp\left(\frac{-(\mathrm{x}-\mu)^2}{2\sigma_{\mathrm{H}}^2}\right)$$

- We first sum over all the probability distributions of the descriptor that correspond to a particular exchange current value $\mathrm{i_0}$. This is expressed as:

$$\hat{\mathrm{p}}(\mathrm{i_0}) = \int_{-\infty}^{+\infty} \mathrm{p_x}(\mathrm{x})\delta(\mathrm{f}(\mathrm{x}) - \mathrm{i_0})\mathrm{dx}$$

The dirac delta function ensures that we sum over all the descriptors value that correspond to a given exchange current value. This needs to be done for every value in the gaussian distribution x. The integral accounts for this complete distribution of the descriptor.

- The probability distribution of the exchange current can now be found by normalizing $\hat{\mathrm{p}}$

$$\mathrm{p_{i_0}}(\mathrm{i_0}) = \frac{\hat{\mathrm{p}}(\mathrm{i_0})}{\int_{-\infty}^{\mathrm{i_{0max}}} \hat{\mathrm{p}}(\mathrm{i_0})}$$

- Expectation value, defined as a probability weighted average, is now obtained using the probability distribution of exchange current. The expectation value of the exchange current can be expressed as:

$$E(i_0) = \int_{-\infty}^{i_{0\max}} i_0 p_{i_0}(i_0)\, di_0$$

The probabilistic activity volcano and the expectation value for exchange current for transfer coefficient (α) of 0.5 can be found in the main text and for transfer coefficient = 1 is shown in the figure S2.

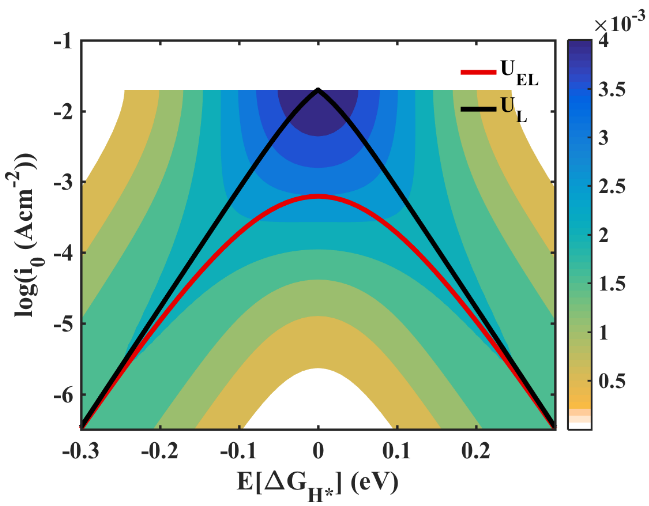

Figure S2: Probabilistic activity volcano for hydrogen evolution with transfer coefficient of 1. The solid black line represents the theoretical exchange current defined by the kinetic model and the red line represents the expectation value of exchange current.

In all the other reaction discussed further, the theoretical activity volcano is constructed on a thermodynamic analysis and not kinetic analysis. Hence in order to compare the prediction efficiency for Hydrogen evolution reaction with the other reactions, we construct a thermodynamic activity volcano and plot the limiting potential as a function of the adsorption free energy of hydrogen on the surface with its associated expectation value in figure S3

4 Chlorine Evolution Reaction

4.1 Reaction Mechanism

We use the Volmer-Heyvrosky reaction mechanism which occurs as follows:
$Cl^-(aq.) + * \rightarrow Cl^* + e^-$

$Cl^* + e^- + Cl^-(aq.) \rightarrow Cl_2(g) + 2e^- + *$

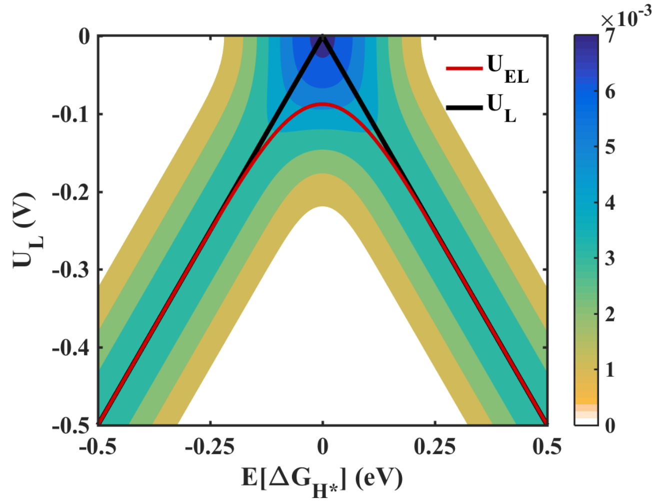

Figure S3: Thermodynamic activity volcano for Hydrogen evolution reaction showing the limiting potential for the reaction as a function of the adsorption free energy of hydrogen. The black line shows the theoretical limiting potential and the red line shows the expectation value of this limiting potential as a function of ΔG_{H^*}

4.2 Calculation Details

The calculations were done on a periodically repeated 4 layered slab for the rutile (110) surfaces of IrO_2, RuO_2, PtO_2, and TiO_2. We consider a 2×1 surface unit cell and $4 \times 4 \times 1$ k-point grid. The bottom two layers were fixed and the top two layers with the adsorbates are allowed to relax. All the structures were converged with a force criterion <0.05 eV$/\mathring{A}$. The surface of the unit cell contains two bridge and two cus sites. Adsorbates bind strongly on the bridge site than on the cus site and therefore the bridge site is occupied with oxygen. Hence, we only focus on the cus sites. We use a 1/2 monolayer coverage (with respect to only the active cus site) of the intermediate on the surface. The adsorption free energy of chlorine is calculated as:

$$\Delta E_{Cl^*} = E(Cl^*) - E(*) - \frac{1}{2}E(Cl_2)$$

$$\Delta G_{Cl^*} = \Delta E_{Cl^*} + \Delta ZPE - T\Delta S = \Delta E_{Cl^*} + 0.37 \text{ (eV)}$$

Figure S5 shows the combined adsorption energy distribution for H* calculated on IrO_2, RuO_2, PtO_2, and TiO_2 for (110) facet. IrO_2 is chosen as the reference because it minimizes the uncertainty in the descriptor (σ_{Cl}).

4.3 Uncertainty Propagation

The methodology used for uncertainty propagation in this case is very similar to the one described for hydrogen evolution.

- The re-centered distribution of adsorption energies of Cl^* is used to find the overall uncertainty in the descriptor (σ_{Cl}). Each of the calculated descriptor values (ΔG_{Cl^*}) is

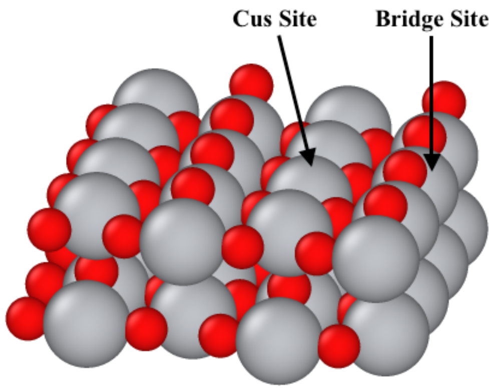

Figure S4: Visualization of the surface structure of rutile oxides (110 facet). Red and grey atoms represent oxygen and metal respectively. Bridges are inactive sites and are occupied with oxygen while the cus sites are the active sites

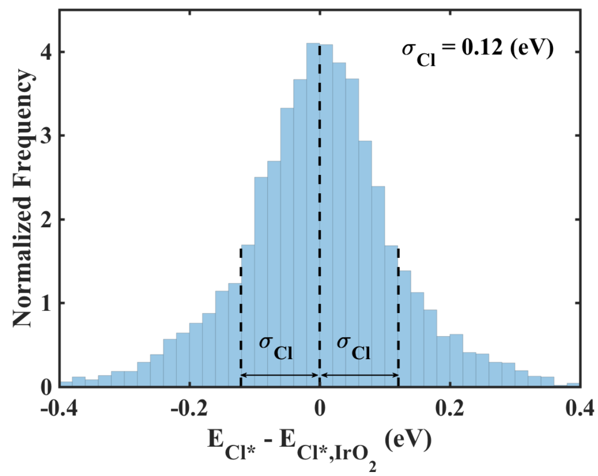

Figure S5: Plot of Normalized frequency as a function of the adsorption energy of the intermediate Cl* relative to IrO_2. The standard deviation of this combined ensemble is σ_{Cl} = 0.12 (eV).

assumed to be a normal distribution with the descriptor value as its mean $\mu = \Delta G_{\mathrm{Cl}^*}$ and standard deviation of σ_{Cl} which can be represented as:

$$X \sim \mathcal{N}(\mu, \sigma_{Cl}^2)$$

- Gaussian distribution is used to find the probability distribution of the descriptor ΔG_{Cl^*}:

$$\mathrm{p_x}(\mathrm{x}|\mu, \sigma_{\mathrm{Cl}}^2) = \frac{1}{\sqrt{(2\pi\sigma_{\mathrm{Cl}}^2)}} \exp\left(\frac{-(\mathrm{x}-\mu)^2}{2\sigma_{\mathrm{Cl}}^2}\right)$$

- For a given value of limiting potential, we sum over all the probability distributions of the descriptors that map to that value of the liming potential. This is expressed as:

$$\hat{p}(U_L) = \int_{-\infty}^{+\infty} p_x \delta(f(x) - U_L)\, dx$$

- Normalized $\hat{p}$ defines the probability distribution of the liming potential as is expressed as:

$$p(U_L) = \frac{\hat{p}(U_L)}{\int_{-\infty}^{U_{L_{max}}} \hat{p}(U_L)}$$

- To calculate the expectation value of the limiting potential, a probability weighted average of limiting potential is calculated.

$$U_{EL} = \int_{-\infty}^{U_{L_{max}}} U_L\ p(U_L)\, dU_L$$

5 Oxygen Reduction Reaction

5.1 4e$^-$ Reaction Mechanism

We consider the following 4e$^-$ associative mechanism for oxygen reduction reaction involving addition of a proton and an electron in each process.

$O_2 + 4H^+ + 4e^- + * \rightarrow OOH^* + 3H^+ + 3e^-$
$OOH^* + 3H^+ + 3e^- \rightarrow O^* + H_2O + 2H^+ + 2e^-$
$O^* + H_2O + 2H^+ + 2e^- \rightarrow OH^* + H_2O + H^+ + e^-$
$OH^* + H_2O + H^+ + e^- \rightarrow 2H_2O + *$

5.2 2e$^-$ Reaction Mechanism

Apart from the 4e$^-$ process, oxygen can also be reduced though a 2e$^-$ pathway in which the single intermediate, OOH*, reduces to hydrogen peroxide H_2O_2. The associative 2^- oxygen reduction proceeds as follows:

$O_2 + 2H^+ + 2e^- +^* \rightarrow OOH^* + H^+ + e^-$
$OOH^* + H^+ + e^- \rightarrow H_2O_2 +^*$

5.3 Calculations Details

Intermediates OH* and OOH* are modeled by including an explicit layer of water to account for hydrogen bonding on a 4 layered $\sqrt{3} \times \sqrt{3}$ configuration for metals and $2\sqrt{3} \times 2\sqrt{3}$ configuration for $Pt_3Ni(111)$ with 1/3 monolayer (ML) coverage. O* is modeled on a 4 layered 2×2 configuration for metals and 2×3 configuration for $Pt_3Ni(111)$ in a fcc site with a 1/4 monolayer (ML) coverage. A $6 \times 6 \times 1$ k-point grid was used for the $2 \times 2 \times 4$ unit cell and the k-points are scaled according to the different unit cells used. The bottom two layers were fixed and the top two layers with the adsorbates are allowed to relax with a force criterion of < 0.05 eV/$\AA$. The adsorption energies od the various intermediates was calculated using the following equations:

$\Delta E_{O*} = E(O^*) - E(^*) - (E(H_2O) - E(H_2))$
$\Delta E_{OH^*} = E(OH^*) - E(^*) - (E(H_2O) - 1/2\ E(H_2))$
$\Delta E_{OOH^*} = E(OOH^*) - E(^*) - (2\ E(H_2O) - 3/2\ E(H_2))$

The entropy corrections and zero point energy corrections can be found in table S1. The gas phase values are from ref. 4 and the values for the adsorbed species are taken from DFT calculations for O and OH adsorbed on Cu(111) from ref. 5 and are assumed to be same for all the metals and the alloy.[4,5] Gas-phase H_2O at 0.035 bar is used as the reference because at this pressure, gas-phase H_2O is in equilibrium with liquid water at 298 K. The entropy correction for the adsorbates on the surface are considered to be zero as the main contribution to the entropy is from the translational entropy.

Table S1: Zero point and entropic corrections at 298 K

	TS	TΔS	ZPE	ΔZPE	$\Delta ZPE - T\Delta S$
$H_2O(l)$	0.67	0	0.56	0	0
$OH^* + 1/2H_2$	0.20	-0.47	0.44	-0.12	0.35
$O^* + H_2$	0.41	-0.27	0.34	-0.22	0.05
$1/2O_2 + H_2$	0.73	0.05	0.32	-0.24	-0.29
H_2	0.41		0.27		
$1/2O_2$	0.32		0.05		
O^*	0		0.07		
OH^*	0		0.30		

As discussed in section 2, the uncertainty in the adoption energy of the various intermediates is found from the combined adsorption energy distribution. The uncertainty in the adsorption energy of O^* is $\sigma_O = 0.21$ (eV); uncertainty in the adsorption energy of OH^* is $\sigma_{OH} = 0.09$ (eV) and uncertainty in the adsorption energy of OOH^* is $\sigma_{OOH} = 0.11$ (eV).

5.4 Uncertainty Propagation

5.4.1 $4e^-$ Reaction Mechanism

For all the metals facets that bond oxygen intermediate too strongly, the limiting potential U_L can be given by:

$$U_L = \Delta G_{OH^*}$$

For the catalysts that bind oxygen intermediate weakly, the limiting potential is given by:

$$U_L = 4.92 - \Delta G_{OOH^*}$$

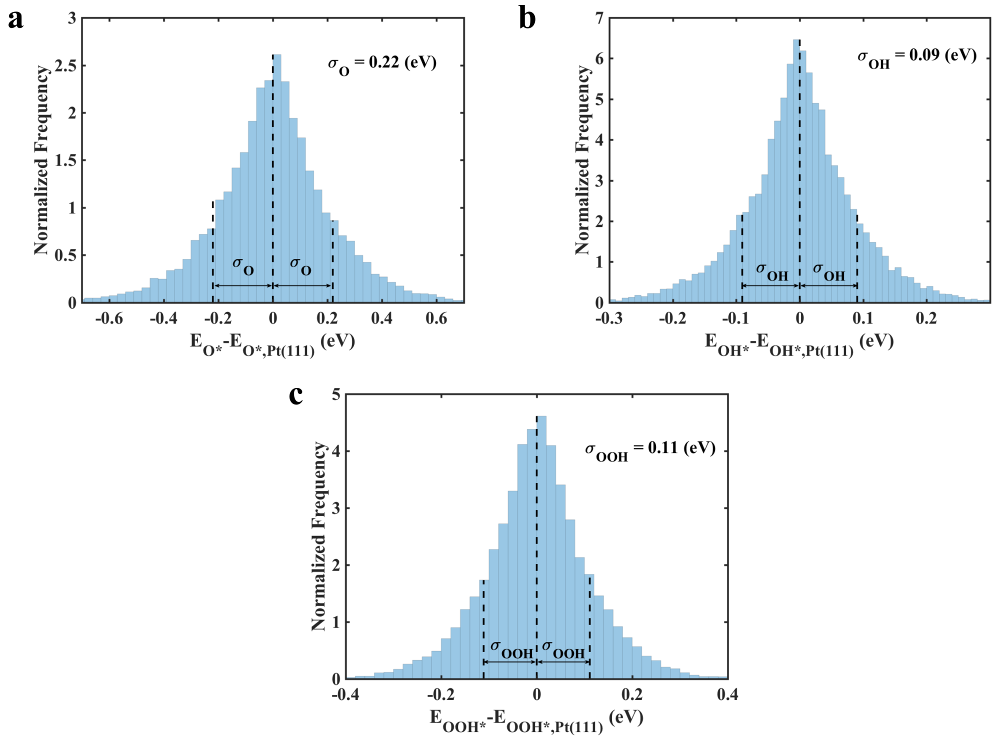

Figure S6: Normalized frequency as a function of the adsorption energy of the intermediate (a) O^*, (b) OH^*, (c) OOH^* relative to Pt(111).

Exploiting the various scaling relations observed for the intermediates involved in the oxygen reduction reaction, we can use ΔG_{O^*},ΔG_{OH^*} and ΔG_{OOH^*} as the descriptor to predict the limiting potential.

(a) Choosing ΔG_{O*} as the descriptor

As the limiting potentials of the stronger binding and weaker binding legs of the volcano are given by ΔG_{OH^*} and ΔG_{OOH^*} respectively, we find a scaling relation between these two quantities as a function of the chosen descriptor, ΔG_{O^*}. The slope for this scaling is fixed to 0.5 and can be rationalized based on the bond conservation principles. The uncertainty in the scaling relation (of the form Y = 0.5*X + C) can be given as:

$$(\sigma_C)^2 = E[(Y - 0.5X)^2] - (E[Y - 0.5X])^2$$

$$(\sigma_C)^2 = E[(0.25X^2 + Y^2 - XY)] - (E[Y] - 0.5E[X])^2$$

$$(\sigma_C)^2 = 0.25E[X^2] + E[Y^2] - E[XY] - 0.25(E[X])^2 - (E[Y])^2 + E[X]E[Y]$$

$$(\sigma_C)^2 = (0.25\sigma_X)^2 + (\sigma_Y)^2 - (E[XY] - E[X]E[Y])$$

$$(\sigma_C)^2 = (0.25\sigma_X)^2 + (\sigma_Y)^2 - (\mu_{XY} - \mu_X\mu_Y)$$

Using the scaling relation shown in figure S7, we can now define the limiting potential in terms of ΔG_{O^*} as follows:

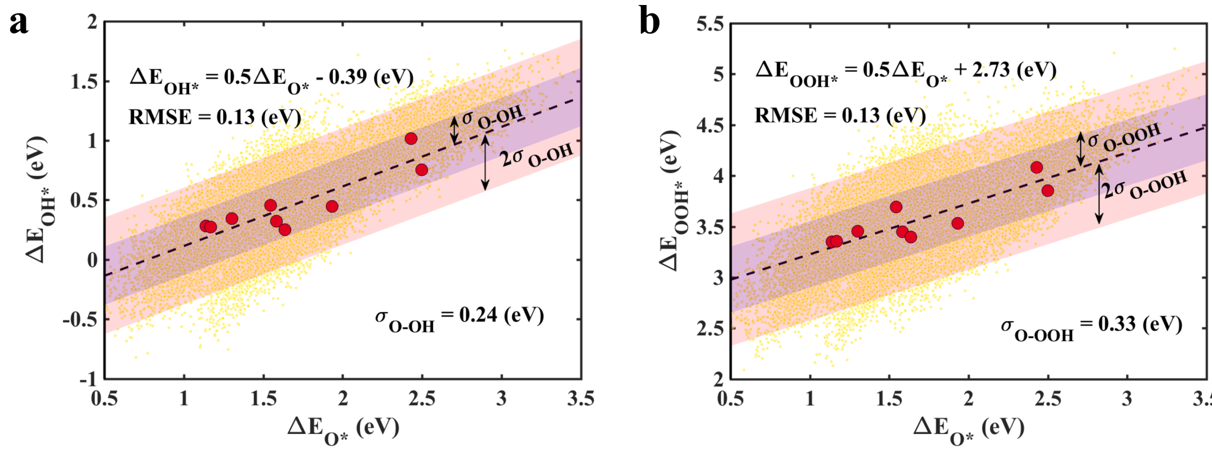

Figure S7: Scaling relation between (a) Adsorption Energy of O^* and OH^* and (b) Adsorption Energy of O^* and OH^*

- For stronger binding leg:

$$\mathrm{U_L} = \mathrm{f}(\Delta \mathrm{G_{O*}}) = \Delta \mathrm{G_{OH^*}} = 0.5\Delta \mathrm{G_{O^*}} - 0.39$$

- For weaker binding leg:

$$\mathrm{U_L} = \mathrm{f}(\Delta \mathrm{G_{O*}}) = 4.92 - \Delta \mathrm{G_{OOH^*}} = 2.19 - 0.5\Delta \mathrm{G_{O^*}}$$

The uncertainty in the stronger binding leg is defined by the uncertainty in the scaling between the adsorption energies of the intermediates O^* and OH^* ($\sigma_{\mathrm{O-OH}}$) and the uncertainty in the weaker binding leg is defined by the uncertainty in the scaling relation between in the adsorption energies of the intermediates O^* and OOH^* ($\sigma_{\mathrm{O-OOH}}$). We use the following methodology to propagate the uncertainty of the descriptor and the scaling relation to the predicted limiting potential:

- Using the uncertainty in the descriptor (σ_{O}) calculated using the re-centered combined distribution of the adsorption energy of O^*, we assume each of the calculated descriptor $\Delta \mathrm{G_{O^*}}$ values as normal distribution with its value as the mean ($\mu = \Delta \mathrm{G_{O^*}}$) and the standard deviation as σ_{O}. This distribution is expressed as:

$$X \sim \mathcal{N}(\mu, \sigma_O^2)$$

- The probability distribution of the descriptor ($\Delta \mathrm{G_{O^*}}$) normal distribution can be found using the Gaussian distribution as:

$$\mathrm{p_x}(\mathrm{x}|\mu, \sigma_{\mathrm{O}}^2) = \frac{1}{\sqrt{(2\pi\sigma_{\mathrm{O}}^2)}} \exp\left(\frac{-(\mathrm{x}-\mu)^2}{2\sigma_{\mathrm{O}}^2}\right)$$

- $\mathrm{U_L}$ function ($\mathrm{f}(\Delta \mathrm{G_{O^*}})$)should take into account the uncertainty in the scaling relation. Hence it can be better represented as:

For stronger binding leg:

$$\mathrm{U_L = f(\Delta G_{O*}}, K_s) = \mathrm{\Delta G_{OH^*} = 0.5\Delta G_{O^*}} + K_s$$

For weaker binding leg:

$$\mathrm{U_L = f(\Delta G_{O*}}, K_w) = 4.92 - \mathrm{\Delta G_{OOH^*}} = K_w - 0.5\mathrm{\Delta G_{O^*}}$$

where, $K_s = \mathcal{N}(\mu = -0.39, \sigma_{\mathrm{O-OH}})$ and $K_w = \mathcal{N}(\mu = 2.19, \sigma_{\mathrm{O-OOH}})$ are normal distributions. Hence for a given value of $\mathrm{\Delta G_{O^*}}$, we have an ensemble of $\mathrm{U_L}$ values which changes the picture of one activity volcano to an ensemble of activity volcanoes with different peaks. For computational purposes we assume this distribution to be discrete and define the random variables $\mathrm{k_s} \in K_s$ and $\mathrm{k_w} \in K_w$. The maximum limiting potential $\mathrm{U_{L_{max}}}$ can be determined by solving the above mentioned two equations simultaneously. For a given activity volcano among the ensemble, at the descriptor value $\mathrm{\Delta G_{O^*} = (k_w - k_s)}$, max limiting potential of $\mathrm{U_{L_{max}} = (k_s + k_w)/2}$ is found. The uncertainty in the descriptor for each of these activity volcanoes is propagated to the limiting potential in a similar manner as described previously and then is averaged for all the activity volcanoes for a given descriptor value.

- Now for an $\mathrm{i^{th}}$ activity volcano relationship in the ensemble we sum over all the probability distribution of the descriptor that map to that value of the limiting potential.

$$\mathrm{\hat{p}_i(U_L) = \int_{-\infty}^{+\infty} p_x \delta(f(x) - U_L)\ dx}$$

- Normalized $\mathrm{(\hat{p})_i}$ defines the probability distribution of the liming potential which is expressed as:

$$\mathrm{p_i(U_L) = \frac{\hat{p}_i(U_L)}{\int_{-\infty}^{(U_{L_{max}})_i} (\hat{p})_i(U_L)}}$$

- To calculate the expectation value of the limiting potential, a probability weighted average of limiting potential is calculated.

$$\mathrm{U_{ELi} = \int_{-\infty}^{(U_{L_{max}})_i} U_L\ p_i(U_L)\ dU_L}$$

- This is done similarly for every member of the activity volcano ensemble and then is averaged over the ensemble for a given $\mathrm{\Delta G_{O^*}}$.

Using this approach, we construct the probabilistic activity volcano with the corresponding expectation value of the limiting potential as shown in the figure S8.

(b) Choosing $\mathrm{\Delta G_{OH*}}$ as the descriptor

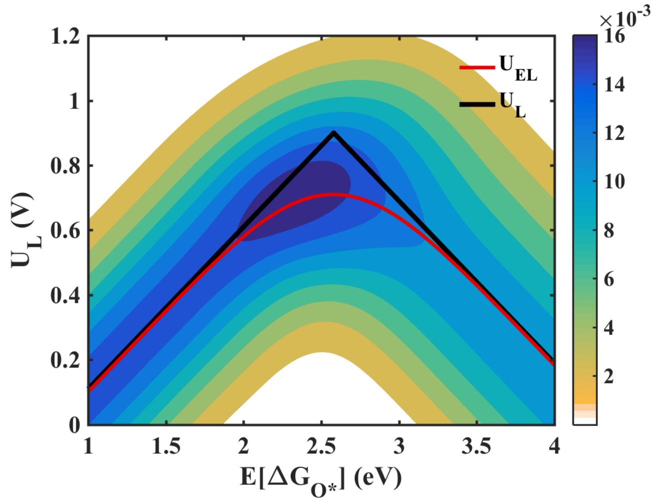

Figure S8: Probabilistic activity volcano for $4e^{-}$ oxygen reduction reaction using ΔG_{O*} as the descriptor. The Prediction Efficiency of this descriptor for the criterion of achieving limiting potential greater than that of Pt(111) is 0%.

The uncertainty in the intercept of the scaling relation (which is of the form $Y = X + C$) between the adsorption energies of intermediate OH^* and OOH^* has been shown in the figure 3(a) can be found in the following way:

$$(\sigma_c)^2 = E[(Y - X)^2] - (E[Y - X])^2$$

$$(\sigma_c)^2 = E[(Y^2 + X^2 - 2XY)] - (E[Y] - E[X])^2$$

$$(\sigma_c)^2 = E[Y^2] + E[X^2] - 2E[XY] - (E[Y])^2 - (E[X])^2 + 2E[X]E[Y]$$

$$(\sigma_c)^2 = (\sigma_Y)^2 + (\sigma_X)^2 - 2(E[XY] - E[X]E[Y])$$

$$(\sigma_c)^2 = (\sigma_X)^2 + (\sigma_Y)^2 - 2(\mu_{XY} - \mu_X\mu_Y)$$

Using the scaling relation between the intermediates OH^* and OOH^* shown in the paper, we can define the limiting potential on terms of the chosen descriptor ΔG_{OH*} as follows:

- For the stronger binding leg:

$$U_L = f(\Delta G_{OH*}) = \Delta G_{OH*}$$

- For the weaker binding leg:

$$U_L = f(\Delta G_{OH*}) = 4.92 - \Delta G_{OOH*} = 1.81 - \Delta G_{OH*}$$

The uncertainty in the stronger binding leg is now defined by the uncertainty in the adsorption energy of the intermediate OH^* (σ_{OH}) and the uncertainty in the weaker binding leg is defined by the uncertainty in the scaling relation between the intermediates OH^* and OOH^* (σ_{OH-OOH}). The following methodology was used to propagate the uncertainty in the descriptor and the uncertainty in the scaling relation to the limiting potential:

- Each of the calculated descriptor value ΔG_{OH*} is assumed as a normal distribution with its value as the mean and the standard deviation given by the uncertainty in the descriptor (σ_{OH}). This normal distribution can be represented as:

$$X \sim \mathcal{N}(\mu, \sigma_{OH})$$

- Using the Gaussian distribution, the probability distribution of descriptor can be found

$$p_x(x|\mu, \sigma_{OH}^2) = \frac{1}{\sqrt{(2\pi\sigma_{OH}^2)}} \exp\left(\frac{-(x-\mu)^2}{2\sigma_{OH}^2}\right)$$

- In order to account for the uncertainty in the scaling relation, the function of limiting potential for the weaker binding leg (defined by scaling) can be better represented as:

$$U_L = f(\Delta G_{OH^*}, K_w) = K_w - \Delta G_{OH^*}$$

  where $K_w = \mathcal{N}(\mu = 1.81, \sigma_{OH-OOH})$ is a normal distribution. Hence a given descriptor value generates an ensemble of predicted limiting potentials, giving rise to an ensemble of activity volcanoes. A discrete distribution of random variables $k_w \in K_w$ is generated to computationally simulate this problem. For a given activity volcano among the ensemble, the maximum limiting potential of $k_w/2$ is found for $\Delta G_{OH*} = k_w/2$. The probability distribution of limiting potential for a given activity volcano among the ensemble is found averaged over the whole ensemble for a given descriptor value as discussed for ΔG_{O^*} as the choice of descriptor.

(c) Choosing ΔG_{OOH*} as the descriptor

The scaling relation relating the adsorption energy of OOH^* and OH^* can be exploited to define the limiting potential in terms of a single descriptor ΔG_{OOH^*}. Hence the limiting potential can be given as:

- For the stronger binding leg:

$$U_L = f(\Delta G_{OOH^*}) = \Delta G_{OH^*} = \Delta G_{OOH^*} - 3.11$$

- For the weaker binding leg:

$$U_L = f(\Delta G_{OOH^*}) = 4.92 - \Delta G_{OOH^*}$$

Hence the uncertainty in the strong binding leg is defined by the uncertainty in the scaling relation and the uncertainty in the weaker binding leg is a function of the uncertainty of the descriptor. To propagate this uncertainty, we follow a similar approach as described for the ΔG_{OH^*} as the descriptor.

- Every calculated value of descriptor is assumed to be normal distribution with the

$\mu = \Delta G_{OOH^*}$ and standard deviation of σ_{OOH^*}. This distribution is expressed as:

$$X = \mathcal{N}(\mu, \sigma_{OOH^*})$$

- The probability distribution of the descriptor can be expressed using the Gaussian distribution as:

$$p_x(x|\mu, \sigma^2_{OOH}) = \frac{1}{\sqrt{(2\pi\sigma^2_{OOH})}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2_{OOH}}\right)$$

- To incorporate the uncertainty in the scaling relation the limiting potential for the stronger binding leg can be represented as:

$$U_L = \Delta G_{OOH^*} - K_s$$

  where $K_s = \mathcal{N}(3.11, \sigma_{OH-OOH})$ is a normal distribution of the scaling intercept. Hence the uncertainty is propagated to an ensemble of activity volcanoes which is then averaged to find the probabilistic activity volcano with the expectation value of the limiting potential.

5.5 $2e^-$ Reaction Mechanism

The metals that bind oxygen intermediates too strongly, removal of OOH^* is the potential determining step and the over potential is given by:

$$U_L = \Delta G_{OOH^*} - \Delta G_{H_2O_2}$$

The activity of the materials binding weakly to the catalysts is associated with the activation of O_2 and the limiting potential is given by:

$$U_L = \Delta G_{O_2} - \Delta G_{OOH^*}$$

Here the formation energies of hydrogen peroxide and oxygen is found using the thermodynamic tables as 3.56 eV and 4.92 eV respectively to avoid the well known issues related to the calculation of molecular reaction energies using DFT. The obvious descriptor for the activity would be ΔG_{OOH^*}, but due to the known scaling between the adsorption energy of OH^*-OOH^* and O^*-OOH^*, ΔG_{OH^*} and ΔG_{O^*} can also be used the descriptors for the limiting potential.

(a) Choosing ΔG_{O^*} as the descriptor

The limiting potentials for both the weaker and stronger binding legs are expressed in terms of the chosen descriptor (ΔG_{O^*}) using the scaling relation as:

- For stronger binding leg:

$$U_L = f(\Delta G_{O^*}) = \Delta G_{OOH^*} - 3.56 = 0.5\Delta G_{O^*} - 0.83$$

- For weaker binding leg:

$$\mathrm{U_L = f(\Delta G_{O^*}) = 4.92 - \Delta G_{OOH^*} = 2.19 - 0.5\Delta G_{O^*}}$$

The uncertainty in the limiting potential is now associated with the uncertainty in the descriptor σ_{O} as well as the uncertainty associated with the scaling relation $\sigma_{\mathrm{O-OOH}}$. In order to propagate the error in both these quantities we follow the uncertainty propagation framework we discussed earlier

- Each of the calculated descriptor values ΔG_{O^*} is assumed to be a normal distribution centered around its value and having a standard deviation of σ_O. This distribution is now represented as:

$$\mathrm{X} \sim \mathcal{N}(\mu, \sigma_{\mathrm{O}}^2)$$

- The probability distribution of X can be obtained using the Gaussian distribution as:

$$\mathrm{p_x(x}|\mu, \sigma_{\mathrm{O}}^2) = \frac{1}{\sqrt{(2\pi\sigma_{\mathrm{O}}^2)}} \exp\left(\frac{-(\mathrm{x} - \mu)^2}{2\sigma_{\mathrm{O}}^2}\right)$$

- In order to account for the uncertainty in the scaling relation, as described earlier, we consider an ensemble of activity volcanoes and hence the limiting potential can be better represented as: For stronger binding leg:

$$\mathrm{U_L = f(\Delta G_{O^*}}, K_s) = 0.5\Delta \mathrm{G_{O^*}} + K_s$$

  For weaker binding leg:

$$\mathrm{U_L = f(\Delta G_{O^*}}, K_w) = K_w - 0.5\Delta \mathrm{G_{O^*}}$$

  where, $K_s = \mathcal{N}(-0.83, \sigma_{\mathrm{O-OOH}})$ and $K_w = \mathcal{N}(2.19, \sigma_{\mathrm{O-OOH}})$ are normal distribution defining the ensemble of activity volcanoes. For each member of the activity volcano, we propagate the uncertainty in the descriptor and find the probability distribution of limiting potential $\mathrm{p(U_L)}$ and the expectation value of the limiting potential as

$$\mathrm{U_{EL}} = \int_{-\infty}^{\mathrm{U_{L_{max}}}} \mathrm{U_L\ p(U_L)\ dU_L}$$

  Unlike the case of $4e^-$ process, where $\mathrm{U_{L_{max}}}$ for each of the activity volcano in the ensemble is determined by the $\mathrm{K_s}$ and $\mathrm{K_w}$ distribution, for $2e^-$ process, the maximum limiting potential is cut off at the equilibrium potential of 0.68 V for all the activity volcanoes. By taking an average over the entire ensemble, we find the overall expectation value of the limiting potential using $\Delta \mathrm{G_{O^*}}$ as the descriptor and taking into account the uncertainty associated with both the descriptor and scaling.

The figure S9 shows the probabilistic activity volcano with the expectation value of the limiting potential for $2e^-$ oxygen reduction reaction using $\Delta \mathrm{G_{O^*}}$ as the descriptor.

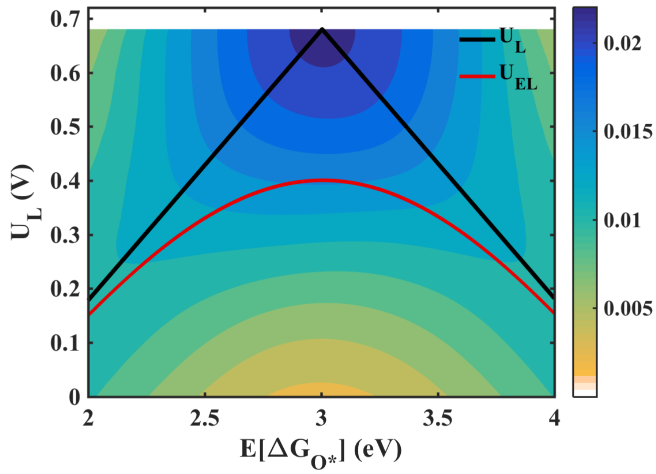

Figure S9: Probabilistic activity volcano for $2e^-$ oxygen reduction reaction using ΔG_{O^*} as the descriptor. The large uncertainty in the descriptor value as well as the scaling relation results in reduced differentiability among materials based on the limiting potential.

(b) Choosing ΔG_{OH^*} as the descriptor

Using the scaling relation between the adsorption energy of the intermediates OH^* and OOH^*, the limiting potential for $2e^-$ ORR can be obtained by the single descriptor ΔG_{OH^*} as follows:

- For the stronger binding leg:

$$U_L = f(\Delta G_{OH^*}) = \Delta G_{OOH^*} - 3.56 = \Delta G_{OH^*} - 0.45$$

- For the weaker binding leg:

$$U_L = f(\Delta G_{OH^*}) = 4.92 - \Delta G_{OOH^*} = 1.81 - \Delta G_{O^*}$$

To propagate the uncertainty in the descriptor σ_{OH} and the uncertainty in the scaling relation σ_{OH-OOH}, we use the same approach described for ΔG_{O^*} descriptor. The uncertainty in the descriptor is propagated by assuming the calculated descriptor value to be a normal distribution and the uncertainty in the scaling is propagated by considering an ensemble of activity volcanoes.

(c) Choosing ΔG_{OOH^*} as the descriptor

Since an activity volcano for $2e^-$ ORR with ΔG_{OOH^*} as the descriptor does not involve scaling, only the uncertainty in the descriptor needs to be propagated to the predicted limiting potential. The uncertainty propagation methodology will therefore be exactly the same as discussed for hydrogen and chlorine evolution reaction which involved a single intermediate in the reaction mechanism. Since there is not uncertainty with respect to scaling that needs to be considered, it is expected that ΔG_{OOH^*} behaves as the most efficient descriptor for $2e^-$.

6 Oxygen Evolution Reaction

6.1 Reaction Mechanism

Oxygen Evolution reaction proceeds with the following associative mechanism:

$H_2O(l) + {}^* \rightleftharpoons OH^* + H^+ + e^-$
$OH^* \rightleftharpoons O^* + H^+ + e^-$
$O^* + H_2O(l) \rightleftharpoons OOH^* + H^+ + e^-$
$OOH^* \rightleftharpoons {}^* + O_2(g) + H^+ + e^-$

The * represents an coordinately unsaturated site (cus) on the rutile oxide (110) surface as shown in S4.

6.2 Calculation Details

The calculations were done on a periodically repeated 4 layered slab for the rutile (110) surface of IrO_2, RuO_2, PtO_2, TiO_2, VO_2, CrO_2, and MnO_2. As discussed in section 4, we consider a 2×1 surface cell and $4 \times 4 \times 1$ k-point grid. We allow the system to relax keeping the bottom 2 layers fixed and allow the top 2 layers with the adsorbates to move. The equations for the adsorption free energy of the intermediates remain same as that discussed in the section 5. Using the methodology discussed in section 2, we find the uncertainty in the adsorption free energy using the combined adsorption energy ensemble. The uncertainty in the adsorption energy of O^* is $\sigma_O = 0.21$ (eV); uncertainty in the adsorption energy of OH^* is $\sigma_{OH} = 0.16$ (eV) and uncertainty in the adsorption energy of OOH^* is $\sigma_{OOH} = 0.16$ (eV).
The ZPE and entropy corrections can be found in table S2. For the adsorbed species the ZPE is obtained from ref. 6 and was calculated for an adsorbate at the cut-site of RuO_2 and is considered to be same for each oxide.[6]

6.3 Uncertainty Propagation

For all the catalysts that bind the intermediates weakly, the limiting potential is given as:

$$U_L = \Delta G_2 = \Delta G_{O^*} - \Delta G_{OH^*}$$

For all the catalysts that bind the intermediates strongly, the limiting potential is expressed as:

$$U_L = \Delta G_3 = \Delta G_{OOH^*} - \Delta G_{O^*}$$

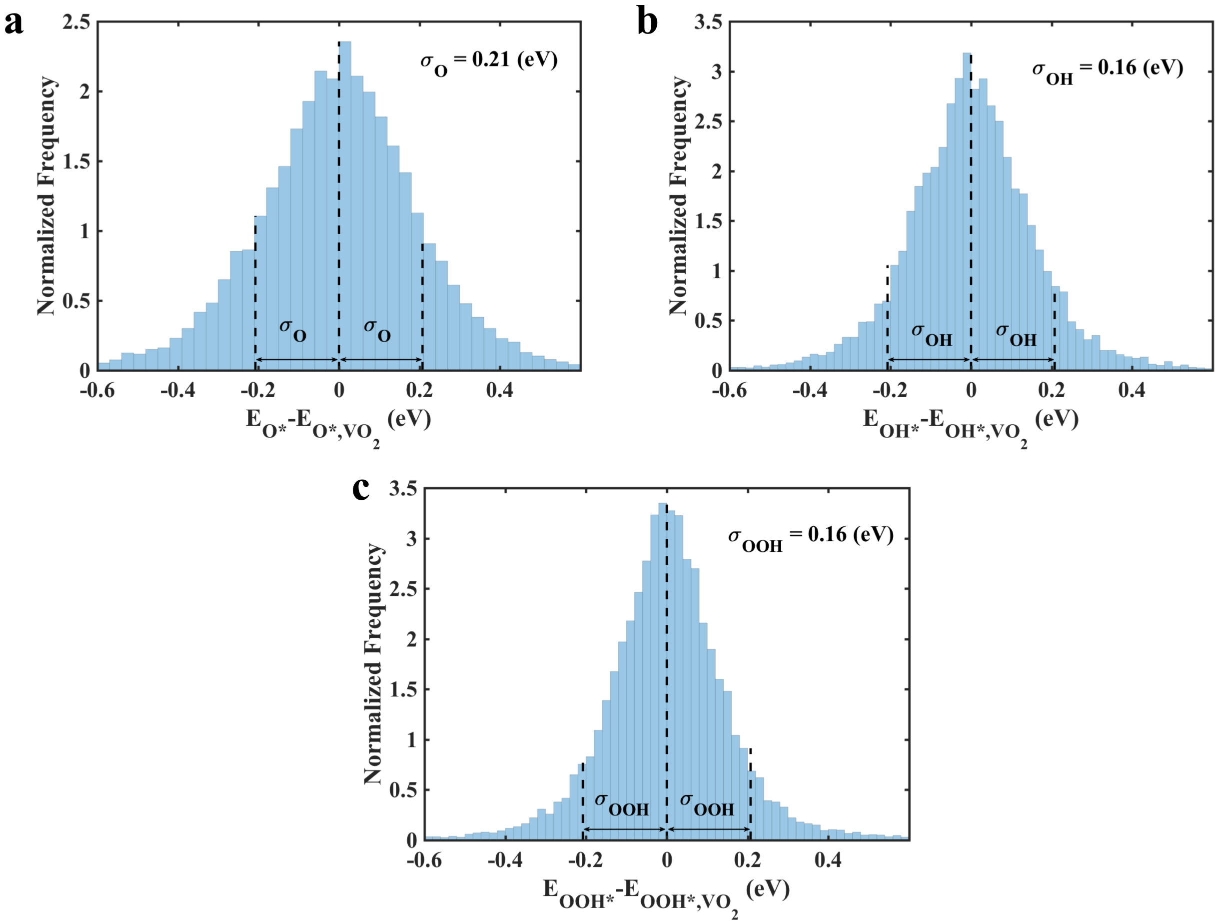

Figure S10: Normalized frequency as a function of the adsorption energy of the intermediate (a) O^*, (b) OH^*, (c) OOH^* relative to VO_2.

Using the scaling relation between the adsorption energies of the intermediates OH^* and OOH^* the limiting potential can be expressed in terms of a unique descriptor: ΔG_2 or ΔG_3. The descriptor that gives a higher prediction efficiency is the one desired to be used.

(a) Choosing $\Delta G_2 = \Delta G_{O^*} - \Delta G_{OH^*}$ as the descriptor

Using the scaling relation we can determine magnitude of the potential determining step (G^{OER}) as:

$$
\begin{aligned}
G^{OER} &= \max[\Delta G_2, \Delta G_3] \\
&= \max[(\Delta G_{O^*} - \Delta G_{OH^*}), (\Delta G_{OOH^*} - \Delta G_{O^*})] \\
&= \max[(\Delta G_{O^*} - \Delta G_{OH^*}), (3.05 - (\Delta G_{O^*} - \Delta G_{OH^*}))] \\
&= \max[\Delta G_2, 3.05 - \Delta G_2]
\end{aligned}
$$

Hence the limiting potential is given as:

Table S2: Zero point and entropic corrections at 298 K for rutile oxides (110)

	TS	TΔS	ZPE	ΔZPE	$\Delta ZPE - T\Delta S$
$H_2O(l)$	0.67	0	0.56	0	0
$OH^* + 1/2H_2$	0.20	-0.47	0.50	-0.06	0.41
$O^* + H_2$	0.41	-0.27	0.34	-0.22	0.05
$1/2O_2 + H_2$	0.73	0.05	0.32	-0.24	-0.29
H_2	0.41		0.27		
$1/2O_2$	0.32		0.05		
O^*	0		0.07		
OH^*	0		0.36		

- For stronger binding leg:

$$\mathrm{U_L} = 3.05 - \Delta \mathrm{G_2}$$

- For weaker binding leg:

$$\mathrm{U_L} = \Delta \mathrm{G_2}$$

Hence the uncertainty in the stronger binding leg is defined by the uncertainty in the scaling relation where as for the weaker binding leg is described by the uncertainty in the descriptor value. The uncertainty propagation method is similar to as described in oxygen reduction reaction with $\Delta \mathrm{G_{OH^*}}$ as the descriptor. Following methodology was used:

- The calculated descriptor value $\Delta \mathrm{G_2}$ is assumed to be normal distribution with the mean given by its value and the standard deviation of $\sigma_{\mathrm{G_2}}$,

$$\sigma_{\mathrm{G_2}} = \sigma_{\mathrm{O}}^2 + \sigma_{\mathrm{OH}}^2 - 2(\mu_{\mathrm{O\times OH}} - \mu_{\mathrm{O}}\mu_{\mathrm{OH}})$$

$$\mathrm{X} \sim \mathcal{N}(\mu, \sigma_{\mathrm{G_2}})$$

- The probability distribution of the descriptor can be expressed using the Gaussian distribution as:

$$\mathrm{p_x}(\mathrm{x}|\mu, \sigma_{\mathrm{G_2}}^2) = \frac{1}{\sqrt{(2\pi\sigma_{\mathrm{G_2}}^2)}} \exp\left(\frac{-(\mathrm{x}-\mu)^2}{2\sigma_{\mathrm{G_2}}^2}\right)$$

- The uncertainty in the scaling is incorporated by considering an ensemble of activity volcanoes as discussed earlier, hence the limiting potential of the stronger binding leg is now expressed as:

$$\mathrm{U_L} = K_s - \Delta \mathrm{G_2}$$

  where $K_s = \mathcal{N}(3.05, \sigma_{\mathrm{OH-OOH}})$ is a normal distribution of the scaling intercept. Uncertainty in each of the activity volcano in the ensemble is propagated in a similar way as described earlier and then averaged over the whole ensemble. The probabilistic activity volcano and the expectation value of the limiting potential obtained from this analysis is shown in the main text.

(b) Choosing $\Delta G_3 = \Delta G_{OOH^*} - \Delta G_{O^*}$ as the descriptor

Using the scaling relationship, we can determine the magnitude of the potential determining step (G^{OER}) in term of ΔG_3 as follows:

$$
\begin{aligned}
G^{OER} &= \max[\Delta G_2, \Delta G_3] \\
&= \max[(\Delta G_{O^*} - \Delta G_{OH^*}), (\Delta G_{OOH^*} - \Delta G_{O^*})] \\
&= \max[(\Delta G_{O^*} - (\Delta G_{OOH^*} - 3.05)), (\Delta G_{OOH^*} - \Delta G_{O^*})] \\
&= \max[\Delta 3.05 - G_3, \Delta G_3]
\end{aligned}
$$

Using this, the limiting potential can be predicted using a single descriptor ΔG_3 :

- For stronger binding leg:
$$U_L = \Delta G_3$$
- For weaker binding leg:
$$U_L = 3.05 - \Delta G_3$$

The uncertainty is propagated by assuming the descriptor value to be a normal distribution value with the mean given by it s value and the standard deviation of σ_{G_3}. This is represented as $X \sim \mathcal{N}(\mu, \sigma_{G_2})$. The uncertainty in the scaling is incorporated in a similar way as described earlier by considering an ensemble of activity volcanoes each corresponding to a different scaling intercept. Using this approach we construct the probabilistic activity volcano with ΔG_3 as the activity descriptor and find the expectation value of the limiting potential as shown in S11

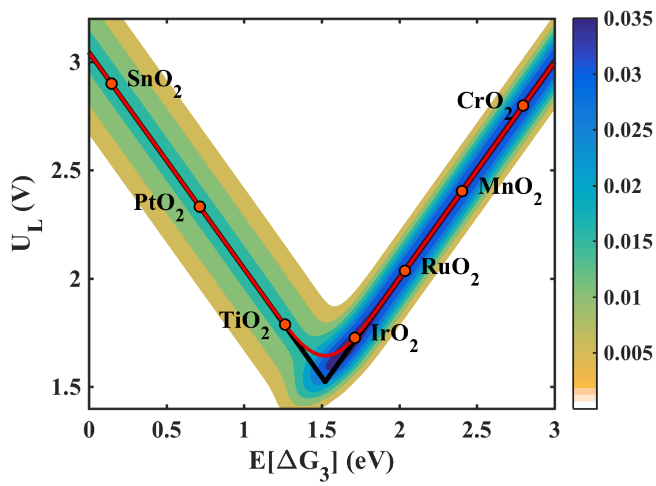

Figure S11: The probabilistic activity volcano for oxygen evolution reaction using $\Delta G_3 = \Delta G_{OOH*} - \Delta G_{O^*}$ as the descriptor.

7 Approaches to Improve Prediction Efficiency

As discussed in the paper, we see prediction efficiency can be improved by using (i) hybrid descriptors (ii) hybrid material references.

We showed earlier that VO_2 and Pt(111) reference minimized the overall uncertainty in the descriptor for oxygen evolution reaction and oxygen reduction reaction respectively. Here we show that using a combination of 2 descriptors for reference enables improving the prediction efficiency. Using a simple space search over the various combinations of the references, we find that a reference of $(0.4\Delta G_{2,TiO_2} + 0.6\Delta G_{2,RuO_2})$ for OER and $0.3\Delta G_{OH^*,Pt(100)} + 0.7\Delta G_{OH^*,Pd(111)}$ for ORR gives higher prediction efficiency.

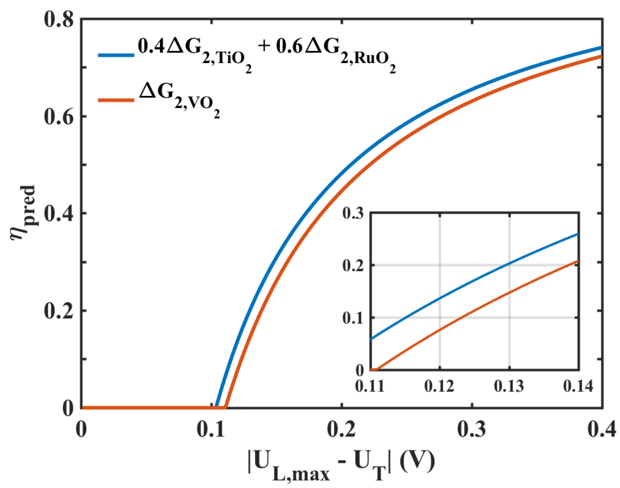

Figure S12: Comparing the prediction efficiency obtained using a single descriptor reference to a hybrid descriptor reference. From the plot it can observed that choosing a hybrid reference of $(0.4\Delta G_{2,TiO_2} + 0.6\Delta G_{2,RuO_2})$ instead of the single reference of $\Delta G_{2,VO_2}$ improved the prediction efficiency and the prediction limit.

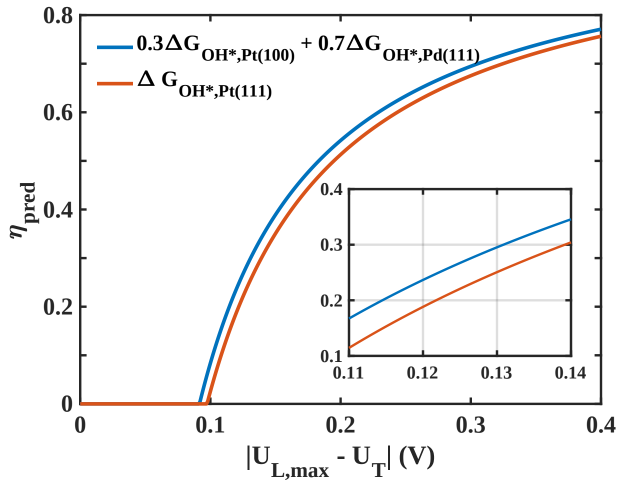

Figure S13: Comparing the prediction efficiency obtained using a single descriptor reference to a hybrid descriptor reference. From the plot it can observed that choosing a hybrid reference of $0.3\Delta G_{OH^*,Pt(100)} + 0.7\Delta G_{OH^*,Pd(111)}$instead of the single reference of $\Delta G_{OH^*,Pt(111)}$ improved the prediction efficiency and the prediction limit.

References

(1) Wellendorff, J.; Lundgaard, K. T.; Møgelhøj, A.; Petzold, V.; Landis, D. D.; Nørskov, J. K.; Bligaard, T.; Jacobsen, K. W. *Phys. Rev. B* **2012**, *85*, 235149.

(2) Deshpande, S.; Kitchin, J. R.; Viswanathan, V. *ACS Catalysis* **2016**, *6*, 5251–5259.

(3) Nørskov, J. K.; Bligaard, T.; Logadottir, A.; Kitchin, J.; Chen, J. G.; Pandelov, S.; Stimming, U. *Journal of The Electrochemical Society* **2005**, *152*, J23–J26.

(4) Atkins, P. *Physical Chemistry*, 6th ed.; Oxford University Press, 1998; pp 485, 925-927, 942.

(5) Kandoi, S.; Gokhale, A.; Grabow, L.; Dumesic, J.; Mavrikakis, M. *Catalysis Letters* **2004**, *93*, 93–100.

(6) Man, I. C.; Su, H.-Y.; Calle-Vallejo, F.; Hansen, H. A.; Martínez, J. I.; Inoglu, N. G.; Kitchin, J.; Jaramillo, T. F.; Nørskov, J. K.; Rossmeisl, J. *ChemCatChem* **2011**, *3*, 1159–1165.