RESEARCH ARTICLE | JUNE 15 2026

Inverse design and discovery of high entropy alloy catalysts for efficient oxygen evolution reaction

Jyotishraj Thoudam ; Ankit Jain

J. Chem. Phys. 164, 234701 (2026)
https://doi.org/10.1063/5.0328756

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Inverse design and discovery of high entropy alloy catalysts for efficient oxygen evolution reaction

Cite as: J. Chem. Phys. 164, 234701 (2026); doi: 10.1063/5.0328756
Submitted: 16 February 2026 • Accepted: 27 May 2026 •
Published Online: 15 June 2026

Jyotishraj Thoudam and Ankit Jain[a)]

AFFILIATIONS

Department of Mechanical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India

[a)]Author to whom correspondence should be addressed: a_jain@iitb.ac.in

ABSTRACT

High-entropy alloys (HEAs) consisting of earth-abundant elements can serve as replacement for expensive, industrial grade catalysts, such as iridium dioxide IrO_2, for oxygen evolution reactions (OERs). However, navigating the exponential configuration space of possible HEAs and unknown inverse structure–property relationship has hindered the development of HEAs for efficient electrocatalytic applications. In this paper, using OH adsorption energy as a descriptor of optimal catalyst property, machine-learning-based frameworks are utilized for discovering HEA catalysts. Each designed HEA is composed of 5 elements chosen from Mn, Fe, Co, Ni, Cu, Zn, and Mo specific to OER. Calculated theoretical overpotential for one HEA catalyst composed of MnCoNiCuZn is found to be $\eta_{theor} = 0.67$ V. However, deviations from scaling results can be as high as 0.99 V when compared to explicit calculations. After screening more than 163×10^9 HEA configurations using a regression model, MnCoNiCuZn with a stoichiometric ratio of 2, 2, 8, 12, and 12 displayed the highest likelihood of positive OH binding energy/active sites. To investigate the intrinsic catalytic activity of the MnCoNiCuZn HEA in an operational OER environment, a computational hydrogen electrode model was constructed to represent the oxyhydroxide layer of OER catalysts, with the MnCoNiCuZn placed at the surface. Analysis indicates that upon consistent comparison, the most stable rutile IrO_2 ($\eta = 0.33$ V) exhibits superior performance relative to the proposed MnCoNiCuZn catalyst ($\eta = 0.57$ V). However, the present methodology, which resulted in statistically favorable adsorption sites, could be successfully used to discover highly active sites within a compositional space.

INTRODUCTION

High-entropy alloys (HEAs) are multi-component alloys, discovered by Cantor *et al.*,[1] which consist of five or more elements, with atomic percentages at 5%–35%.[2] They are preferred over conventional alloys due to their attractive mechanical and energetic properties, which include remarkable mechanical strength, thermal stability, exceptional corrosion resistance, and existence of several adsorption sites due to an exponentially large number of configurations.[3,4] The physico-chemical characteristics of HEAs, such as stability, phase, and structure, can be predicted from the following semi-empirical parameters: entropy of mixing (ΔS_{mix}), enthalpy of mixing (ΔH_{mix}), atomic radius difference factor (δ), and valence electron configurations (VEC).[5] Recently HEAs have been reported[6,7] as a viable candidate for serving as a design platform for discovering promising catalysts for their large unexplored active sites. HEAs composed of earth-abundant metals open up the possibility of designing sustainable,[8,9] scalable,[10] and efficient catalysts that can replace expensive industrial grade oxygen evolution reaction (OER) catalysts such as IrO_2 that are currently used.[11] This is due to the accessibility of different active sites offered by numerous configurations for a given structure and stoichiometric ratio. However, exploring the attractive chemical spaces of HEAs is hindered by three obstacles.[12] First, limited and time consuming experiment to accurately synthesize the sheer magnitude of configurations resulting from exponential arrangements allowed by HEAs for different array of atomic percentage.[13] Second, exploring all possible structure and configurations of HEAs by computational techniques such as density functional theory (DFT) are computationally expensive. Third, inverse-structure–property relationships between atomic configurations and desired energetics remain elusive.

In this paper, new HEA compositions are explored and an optimal HEA is proposed that can serve as catalyst for oxygen evolution reaction using machine learning accelerated DFT simulations. In the first stage, the widely used elements in catalysts are identified specific to OER experiments.[7,11] Each 5 element HEA catalyst is chosen from 7 d-block elements. Combinatorial calculations reveal ${}^{7}C_5 = 21$ different HEA types. Second, the empirical structure–property relationships of HEAs are leveraged for designing solid phase; FCC HEAs and the problem of navigating large configuration space are addressed by choosing FCC lattice.[5] This is followed by generation of the OH adsorption/binding energy ($\Delta E_{ads,OH}$) database using DFT, which will be used as a training dataset for machine learning models. As OH binding energy is a good descriptor for identifying the OER catalytic performance,[14–16] it is chosen as the target property for training the machine learning models. In the third stage, a traditional autoencoder model[17] and a linear regression model are trained to predict the target property for obtaining new HEAs that exhibit positive OH binding energy configurations. Each of the predicted HEA configurations from machine learning models are tested against DFT simulations implemented in VASP.[18] To enhance the performance of the machine learning model, the regression model is iteratively trained for six cycles. During training, some of the detected HEA configurations displayed near optimum OH binding energy ~0.8 eV favorable for OER.[14] These configurations were found to be near the apex of the theoretical overpotential vs the $\Delta G_{O^*} - \Delta G_{OH^*}$ curve. The theoretical overpotential curve was derived using the linear scaling relationship of $\Delta G_{OOH^*} = \Delta G_{OH^*} + 3.2$ eV, in the context of the sluggish four steps OER reaction mechanism known as adsorbate evolution mechanism (AEM). Note that the final machine learning model correctly predicted that more than 50% of HEA configurations exhibit positive OH binding energy. Furthermore, the likelihood of detecting active OER sites for a given stoichiometric ratio for each of the 21 different HEA types were screened based on the percentage of positive $\Delta E_{ads,OH}$. Finally, the catalytic performance for the most promising HEA type is evaluated using a computational hydrogen electrode[19] emulating actual experimental environment and compared with rutile (110) IrO_2.

METHODS

Creation of HEA slabs

Earth-abundant elements that belong in the d-block of the periodic table are chosen for creating HEA crystals. Based on a frequency study of the entire periodic table, the current most widely used elements in electrocatalyst specific are first identified.[11] Manganese (Mn), iron (Fe), cobalt (Co), nickel (Ni), copper (Cu), zinc (Zn), and molybdenum (Mo) are chosen, as shown in Fig. 1(a). It is important to note that earth-abundant metals exhibit a variety of crystal structures, including face-centered cubic (FCC), body-centered cubic (BCC), and hexagonal close-packed (HCP). For example, Ni and Cu adopt FCC structures, Mo is BCC, while Co and Zn are HCP, and Mn exhibits a complex cubic structure.[20–22] Although many constituent elements in high-entropy alloys adopt different crystal structures in their elemental form, FCC solid solutions are stabilized

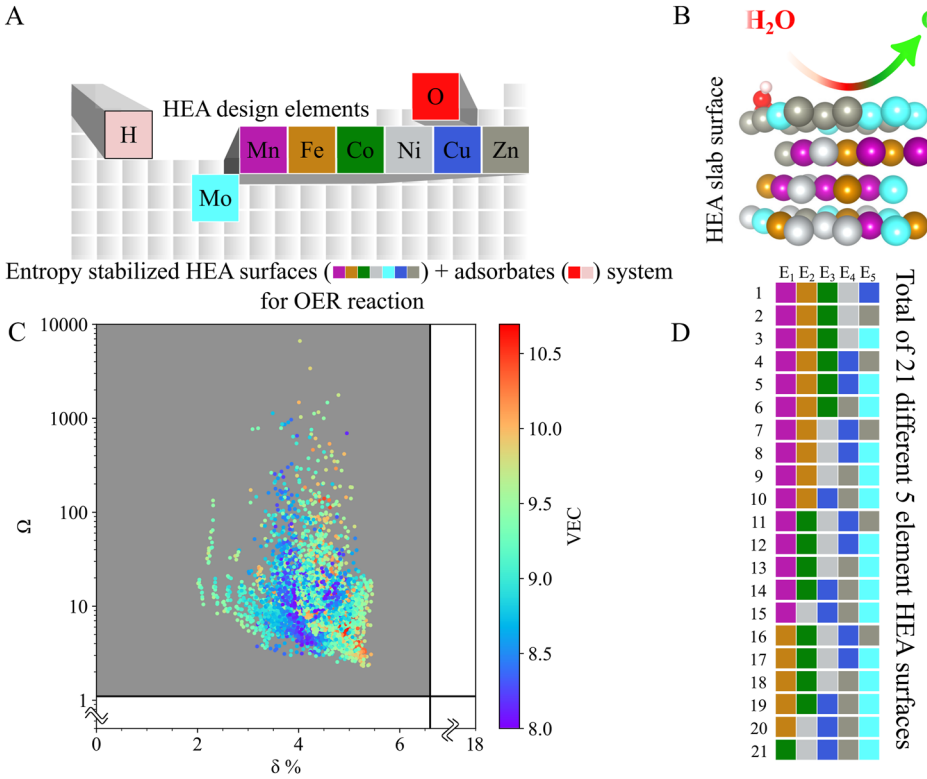

FIG. 1. Creation of HEA solid surfaces for oxygen evolution reaction. (a) Each 5-element HEA crystal designed by choosing from seven d-block elements Mn, Fe, Co, Ni, Cu, Zn, Mo. (b) A sample of generated FCC (111) HEA-high entropy crystal surface for adsorption energy calculations MnFeNiZnMo for oxygen evolution reaction. (c) Ω–δ–VEC values for all the randomly chosen HEA crystals before high throughput calculations fall within the theoretical limits $\Omega \geq 1.1$ and $\delta \leq 6.6\%$ for solid solutions. Empirical conditions for formation of the FCC phase owing to VEC > 8 are also observed to be satisfied.[5,23,24] Computations of ΔH_{mix} uses Miedema's model.[25,26] Note that $\Omega = \frac{T_m \Delta S_{mix}}{|\Delta H_{mix}|}$. (d) A total of 21 different combinations of possible HEA catalyst are shown. To satisfy HEA atomic percentages of 5%–35%, each element (E_1, E_2, E_3, E_4, E_5) must occupy no less than 2 and no more than 12 atoms in the created surface slab of the 36 atom supercell.

and accessible in HEAs due to configurational entropy and other thermophysical properties. HEA crystals for this study are generated in FCC configurations constraining the structure space. Before DFT computations of the potential FCC slab, each of the chosen HEAs are evaluated and found to satisfy the semi-empirical solid solution criterion of Gibbs free energy factor $\Omega\left(=\frac{T_m\Delta S_{mix}}{|\Delta H_{mix}|}\right) \geq 1.1$, which relates the entropic stabilization and enthalpic destabilization ratio of the alloy mixture and the atomic radius difference factor $\delta \leq 6.6\%$, including the condition for the formation of FCC phase where the average valence electron configuration $VEC > 8$.[5,23,24] This is shown in Fig. 1(c). Here, T_m, ΔS_{mix}, and ΔH_{mix} are the melting temperature, mixing entropy, and mixing enthalpy of the 5 element alloy. In order to compute Ω, Miedema's model[25,26] was used to calculate all pairwise mixing enthalpy (H_{ij}), which is a necessary input for computing $\Delta H_{mix} = \sum_{i=1,i<j}^{5} 4H_{ij}c_ic_j$. Further details are provided in the supplementary material for the rest of the parameters. Constraining the crystal structure reduces the space for computational navigation of exponentially large HEA configurations offering tractability.

After fixing the HEA lattice to FCC, HEA slabs are generated using ASE (atomic simulation environment)[22] to create a database along with adsorption energies of hydroxide ion (OH) as the corresponding energy target descriptor for machine learning. Each of the generated HEA slabs are composed of five different elements chosen from Mn, Fe, Co, Ni, Cu, Zn, and Mo. Slabs consisting of four layers with 36 atoms in their lattice sites are created with supercell size of 3 × 3 × 4 times the unit cell along the (111) direction.

To ensure the atomic compositions satisfy HEA percentages of 5%–35%,[2] the five elements must populate a minimum of 2 and a maximum of 12 lattice sites. Typically, the lattice parameter of an HEA is computed using percentage weighted lattice constant of the respective bulk crystals of the elements present in the HEA slab.[6] However, most of the elements considered in this study, such as manganese (Mn), do not exhibit cubic FCC structures. Therefore, the lattice parameter for generating an FCC HEA slab containing 36 atoms is extracted from their respective bulk unit cells using a Voronoi volume tessellation. A sample HEA slab is shown in Fig. 1(b). Density functional theory implemented in VASP with spin-polarization with Grimme's D3 correction[27] is used for all DFT energy calculations. Periodic boundary condition is imposed on all directions, and a vacuum spacing of 10 Å is imposed on both sides of the surface employing a k-point grid size of 3 × 3 × 1 for all adsorption energy computations. The energy cutoff is taken to be 400 eV. PAW pseudopotential[28] with PBEsol exchange–correlation functional[29] is used where convergence criterion for force is 0.05 eV/Å. The bottom two layers are fixed and the top two upper layers are allowed to relax during optimization of geometry.

After identifying the optimized HEA catalyst from machine learning, DFT+U calculations[30] are carried out to evaluate the catalytic performance in actual experimental environment using multiple initial magnetic moments.

Extraction of OH adsorption energy ($\Delta E_{ads,OH}$)

Adsorption energies for hydroxyl molecule (OH) were extracted for 21 000 HEA slabs for building the initial training database for machine learning (ML) to predict HEA configurations for superior catalytic performance for OER.[14–16] Adsorption free energies (ΔG) are taken into account at a later stage by adding zero-point energies (ZPE) and temperature-dependent enthalpy of vibrations (H_{vib}) for each adsorbate and liquid water molecule.[31] Since there are 21 different, shown in Fig. 1(d), ways of choosing 5 elements from a set of 7 elements, 1000 random atomic percentages are chosen for each 5-element set. OH adsorption energies were computed using the following equation for relaxed geometries:[6]

$$\Delta E_{ads,OH} = E_{slab+OH^*} - E_{slab} + \left(\frac{1}{2}E_{H_2} - E_{H_2O}\right), \quad (1)$$

where $E_{slab+OH^*}$ is the DFT electronic energy of the slab with adsorbed OH, E_{slab} is the electronic energy of the HEA slab, E_{H_2} is the electronic energy of H_2 molecule, and E_{H_2O} is the electronic energy of the H_2O molecule. The molecular energies E_{H_2} and E_{H_2O} were computed with the same settings as the slabs, but with 10 × 10 × 10 Å^3 unit cells and k-points grid size of 1 × 1 × 1.

Representation of a slab with an adsorbate for machine learning

Performance of ML models depends upon the representation of the dataset. To represent each of the HEA slabs uniquely, the seven chemical elements (Mn, Fe, Co, Ni, Cu, Zn, and Mo) are one-hot encoded. Since each slab crystal contains 36 atoms, which are arranged in a 3 × 3 array consisting of 9 atoms in 4 layers, and the shape of the one-hot encoded tensor is chosen to be 4 × 3 × 3 × 7. The adsorbate (OH) molecule is also modeled as one-hot encoded 2D matrix, with shape 3 × 3. This is modeled based on the adsorption sites available to the OH molecule, which is near one of the 9 atoms in the top layer. The values of the target property, $\Delta E_{ads,OH}$, are scaled using the mean and standard deviation based on the training set.

RESULTS AND DISCUSSION

Developing a robust high-throughput machine learning model

The adsorbate OH molecules were placed slightly perturbed from the on-top location of the HEA slabs. The initial OH binding/adsorption energy distribution computed using density functional theory can be seen on the light blue histogram of Fig. 2(d) labeled as "Before Machine Learning." Among the 21 000 HEA slabs with random configuration, less than 2.4% exhibited positive OH binding energy. Since the supercell consists of 36 atoms more than ~10^{21}, HEA configurations are possible for one stoichiometric ratio. Therefore, to navigate through this vast configuration space, a traditional autoencoder model was employed to predict the desired configuration for a given OH binding energy. Note that even though the OH molecule is placed at a perturbed on-top site in the initial geometry, the molecule is found to relax to a hollow site in 80% of the configurations (tested for 25 randomly selected configurations from MnCoNiCuZn composition). Furthermore, in more than 10% of the analyzed configurations, OH relaxed in the horizontal orientation, although the initial orientation was almost vertical for all configurations. In addition to the traditional framework of encoder–latent space–decoder model,[32] a feed forward regressor was attached to the latent space, as shown in Fig. 2(a). The one hot-encoded HEA

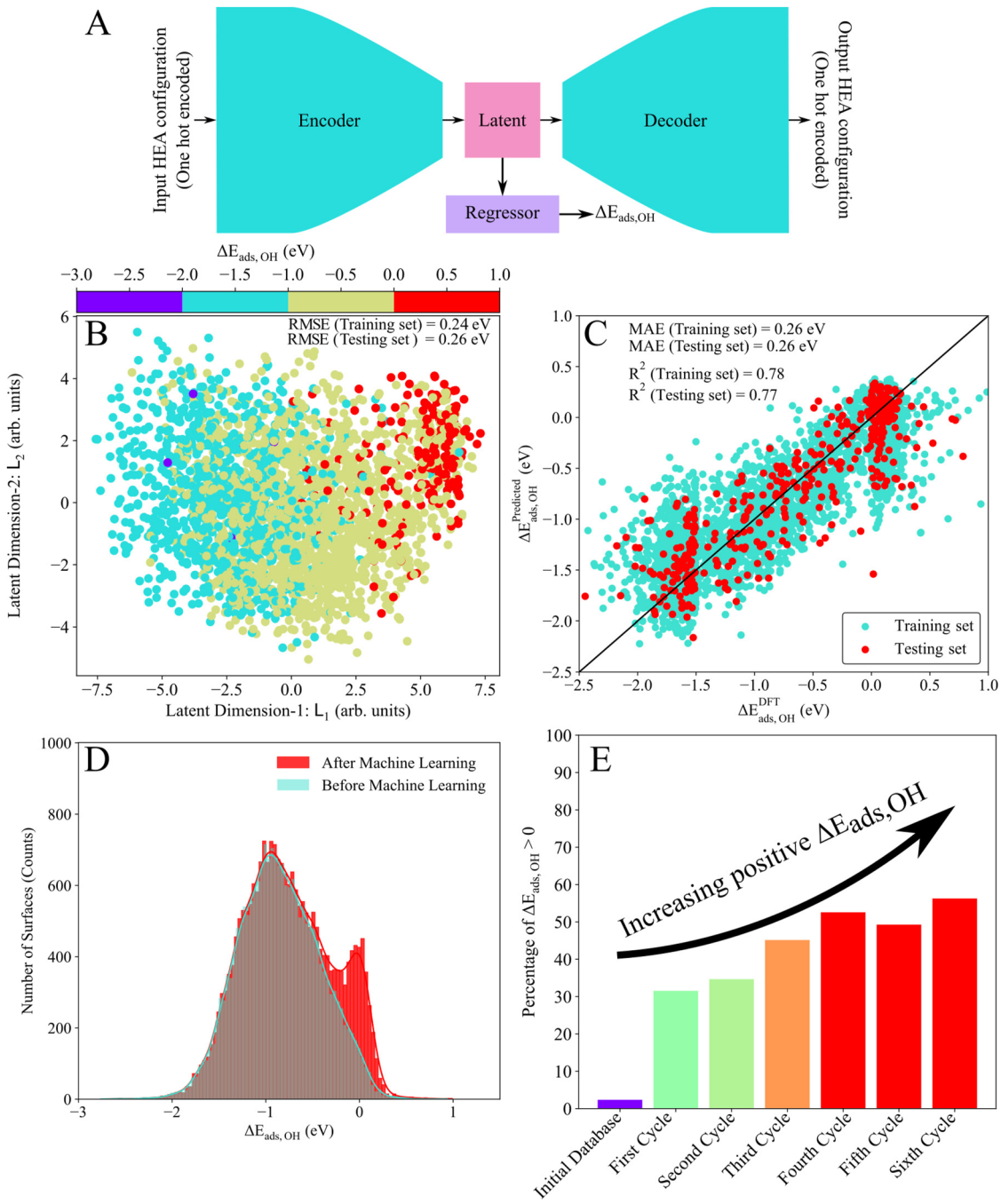

FIG. 2. High-throughput machine learning assisted search for desirable HEA configurations that contains active sites with positive OH binding energies ($\Delta E_{ads,OH} > 0$ eV) validated by density functional theory calculation. Initially generated database of $\Delta E_{ads,OH}$ for ~21 000 HEA slabs using density functional theory implemented in VASP displayed less than 3% of slab configuration possessing positive OH binding energies. This database is used as input to (a) A traditional autoencoder architecture, with a regressor attached to the latent space for predicting $\Delta E_{ads,OH}$ and to learn desirable HEA configurations. (b) Distinct clustering of HEA configurations corresponding to OH binding energy is seen. However, the error in the predicted energies is ~0.26 eV, which is similar to that in the (c) linear regression model. (d) Therefore, linear regression is employed for faster sampling of configurations. Contrasting counts of desirable HEA configurations before and after machine learning on $\Delta E_{ads,OH}$ and the final data distribution is presented. (e) Increasing trend of positive $\Delta E_{ads,OH}$ configurations was detected after each learning cycle. A total of six cycles were performed. Detection of desirable HEA configurations increased up to more than 50%.

tensor along with the 2D adsorbate layer, as described above, was used as the feature vector, with OH binding as the descriptor. The autoencoder clusters the HEA configurations in the latent space based on OH binding energies while minimizing the error during training. In Fig. 2(b), the red circles are clustered together, denoting the positive OH binding energy configurations. For autoencoders, uncertainties associated with choosing different ways of combining latent space vectors presents difficulties in generating new HEA configurations by passing new latent space points through the decoder. It was observed that the error for the prediction of novel configurations that resulted in positive OH binding energy was 0.26 eV. Surprisingly, using a simple linear regression model, the same MAE error was observed as shown in Figs. 2(b) and 2(c). The details of machine learning frameworks are provided in the supplementary material. In addition, the prediction of novel HEA configurations that showed positive OH binding energy was higher than that of traditional autoencoder models. The linear regression model was first trained upon the initial 21 000 $\Delta E_{ads,\mathrm{OH}}$ database. Using the trained model, a new set of 500 HEA configurations exhibiting positive $\Delta E_{ads,\mathrm{OH}}$ was predicted. After validation of the predicted configurations using DFT, the percentage of positive OH binding energy configurations are plotted for each 6 cycles in Fig. 2(e). After six cycles, the final DFT database consisted of ~23 000 HEA slabs with $\Delta E_{ads,\mathrm{OH}}$ values, which is shown in Fig. 2. In Fig. 2(e), it was observed that the model was able to perform better at detecting active catalysts compared to random search. The model improves from detecting 31% of positive OH binding energy HEA configurations in the first iteration of upto 56% in its sixth iteration. Since the linear regression model is extremely fast in predicting configurations due to its relative simplicity compared to the autoencoder model, it is used for further high-throughput screening of HEA configurations. It is also important to note that each HEA system is represented by a total array length of more than 252 ($4 \times 3 \times 3 \times 7$), which is positionally encoded. The positional encoding is designed, in part, to capture the local chemical environment. However, more expressive representations of the crystals have also been explored using graph neural networks and SOAP-GAP models. In particular, a MACE model[33] fine-tuned on the existing "mace-mp-0b3-medium.model" and SOAP-GAP potential[34,35] is trained using the energies, forces, and configurations from the ionic relaxation trajectories of the proposed MnCoNiCuZn catalyst. The details of the models are added in the supplementary material. Using these models, the adsorption energies are computed. The fine-tuned MACE model achieves an MAE of 0.10 eV, while the SOAP-GAP model performed poorly (MAE of 0.58 eV). The MACE model (0.10 eV) outperforms the linear regression model (0.26 eV). However, the latter remains advantageous for high-throughput screening applications. In particular, for one HEA configuration, the MACE model requires ~60 s per inference, whereas the linear regression model operates on a timescale of ~10^{-5} seconds, allowing for efficient evaluation of vast configurational spaces (~163×10^9) in this case. The details of the training can be found from the supplementary material.

Exploring exponential configuration space of HEA catalysts

During the iterative cycle of training the linear regression model, we observed the elemental HEA types that exhibited OH binding energy of more than 0.5 eV. The highest number of detected HEA types that displayed more than this $\Delta E_{ads,\mathrm{OH}}$ threshold were those of MnCoNiCuZn. The linear adsorption energy scaling relationship developed by Man *et al.*,[14] which is preserved in adsorption free energy[36] ($\Delta G_{\mathrm{OOH}^*} = \Delta G_{\mathrm{OH}^*} + 3.2$ eV) due to negligible zero-point energies, is employed along with the widely used relationship,

$$\Delta G_{O^*} = 2 \times \Delta G_{\mathrm{OH}^*} \tag{2}$$

without an intercept[36] to derive the theoretical overpotential, $\eta_{theor}(V)$. In this calculation, $\Delta E_{ads,\mathrm{OH}}$ is used to compute the adsorption free energy of OH (ΔG_{OH^*}) using the following relation:[31]

$$\begin{aligned}\Delta G_{\mathrm{OH}^*} = {} & \Delta E_{ads,\mathrm{OH}} + \Delta ZPE_{\mathrm{OH}^*} + \Delta H_{vib,\mathrm{OH}^*} - \Delta TS_{\mathrm{OH}^*} \\ & - \{\Delta ZPE_{\mathrm{H_2O}} + \Delta H_{vib,\mathrm{H_2O}} - \Delta TS_{\mathrm{H_2O}} \\ & - 0.5 \times (\Delta ZPE_{\mathrm{H_2}} + \Delta H_{vib,\mathrm{H_2}} - \Delta TS_{\mathrm{H_2}})\}.\end{aligned} \tag{3}$$

The *ZPE*, H_{vib}, and *TS* data are taken from the work of Zhang *et al.*[31] where the temperature is 300 K. Using Eq. (2), linear scaling relationship, and Eq. (3) in the four concerted proton-electron transfer AEM oxygen evolution reaction steps, the theoretical minimum overpotential is 0.37 V at typical OER descriptor, $x = \Delta G_{O^*} - \Delta G_{\mathrm{OH}^*} = 1.6$ eV.

The x value is recomputed using the above-mentioned scaling relations and theoretical overpotential is evaluated for the predicted catalysts found during the iterative cycle, which exhibited OH binding energies greater than 0.5 eV. Using scaling relations, the HEA-type MnCoNiCuZn denoted as "Catalyst-1," shown in Fig. 3(a), resulted in the lowest overpotential η_{theor} of 0.67 V. Compared to literature reported overpotentials of 0.61–0.67 V[37,38] for IrO_2, an overpotential of 0.33 V for IrO_2 was obtained using the computational settings discussed in the supplementary material, which are consistent with other calculations reported in this work. However, literature studies suggest that simple scaling relations may not hold true for the considered HEAs. Therefore, the overpotentials of the selected nine catalysts was also computed using the computational hydrogen electrode with explicit calculations of all reaction intermediate energies, which allows us to test the validity of scaling relations for these systems. It was found that the computed overpotentials were between 1.11 and 1.79 V, which are higher than the ones predicted by the scaling relations (0.67–1.14 V). The full list of overpotentials obtained for considered nine catalysts, along with differences in adsorption energies of reaction intermediates are listed in Table S5 in the supplementary material. Furthermore, the actual value $\Delta = \Delta G_{\mathrm{OOH}^*} - (\Delta G_{\mathrm{OH}^*} + 3.2)$ was investigated for the nine metallic HEA catalysts presented in this work using a computational hydrogen electrode. It was found that Δ varied from 0.16 to 0.59 eV. Data suggest that among the nine catalysts, seven of them deviated from the universal scaling relationship.

Due to difficulty in predicting the consistency of the result as it may be an outlier configuration among the multitude of possible configuration for that particular stoichiometric ratio, more configurational space are explored further. Therefore, part of the exponential configuration space, more than 163×10^9 configurations, resulting from combined stoichiometric ratios for all HEA types are explored. For each 21 HEA types, there are 8725 possible

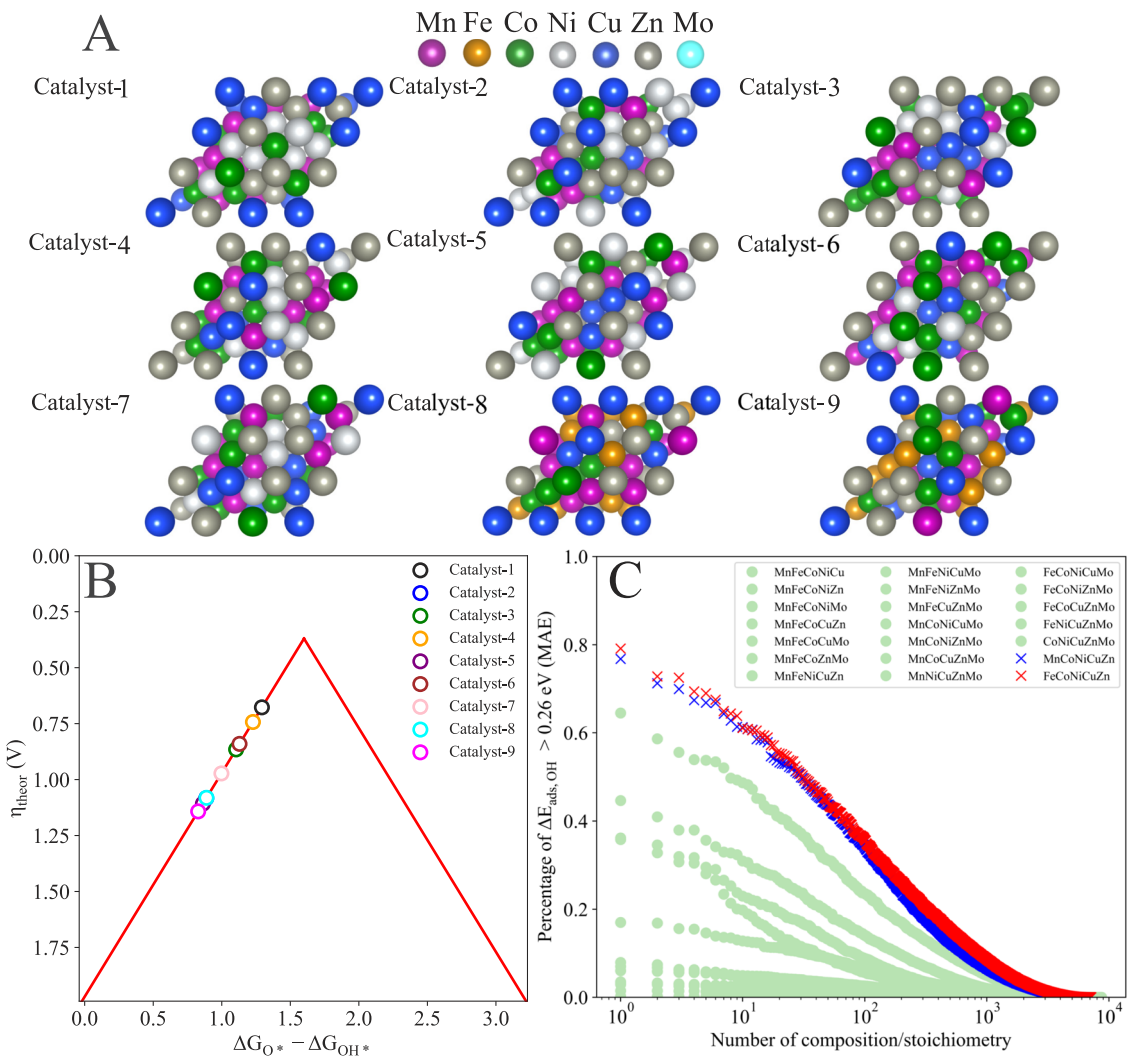

FIG. 3. Theoretical overpotential plot and exploration of HEA configuration space to find the desired stoichiometric ratio that gives the highest positive OH binding energy likelihood using the trained linear regression model. (a) Nine HEA catalyst surfaces with $\Delta E_{ads,OH} > 0.5$ eV lying near the apex of the volcano curve, detected by the linear regression model after DFT validation. (b) As linear regression model improves, it started detecting HEA configurations consisting mostly of MnCoNiCuZn that possess near-optimum OH binding energy ($\Delta E_{ads,OH} \sim 0.8$ eV). Therefore, detected HEA configurations with $\Delta E_{ads,OH} > 0.5$ eV are used to compute theoretical overpotential (η_{theor}) derived using the linear scaling relationship $\Delta G_{OOH^*} = \Delta G_{OH^*} + 3.2$ eV. Catalysts-1 to -7 are composed of MnCoNiCuZn, with the lowest overpotential for Catalyst-1 configuration with $\eta_{theor} = 0.67V$, while Catalyst-8 and -9 are composed of MnFeCoCuZn. (c) Percentage of positive $\Delta E_{ads,OH}$ for 1×10^6 configuration for all 21 HEA types at each stoichiometric ratio is predicted using the trained linear-regression model. A total of more than 163×10^9 HEA configurations were tested. FeCoNiCuZn and MnCoNiCuZn with elemental composition Fe/Mn = 2, Co = 2, Ni = 8, Cu = 12, and Zn = 12 was found to possess the highest possible positive OH binding energy configurations of nearly 1% per million configurations. Prediction of MnCoNiCuZn is consisted with those of Catalyst-1 to 7, and therefore, the catalytic performance is investigated further that resembles experimental environment.

stoichiometric ratios. Therefore, for each 21 × 8725 HEA compositions, ~1 000 000 HEA configurations are fed as an input to the linear regression model and OH binding energies are predicted. Afterward, the likelihood of a stoichiometric ratio for each HEA types that exhibits positive OH binding energies is quantified using the following formula:

$$\text{Percentage of } \Delta E_{ads,OH}(> 0.26 \text{ eV}) = \frac{\text{Number of positive HEA configurations}}{1\,000\,000} \times 100. \quad (4)$$

Here, only those HEA configurations with MAE of above 0.26 eV for linear regression are chosen to remove the uncertainties of

the trained model. The percentage likelihood that an HEA type would serve as a good catalyst can be seen in Fig. 3(b). After screening a total of more than 163 × 10^9 HEA configurations, an optimized catalyst with composition was identified. It was found that X–CoNiCuZn (X = Mn/Fe), with stoichiometric ratio X:Co:Ni:Cu:Zn::2:2:8:12:12, exhibited the most fertile ground for oxygen evolution reaction with nearly 1% out of 1 × 10^6 configurations possessing positive $\Delta E_{ads,\mathrm{OH}}$ values. Although $\Delta E_{ads,\mathrm{OH}}$ is useful as a screening descriptor, the final evaluation of OER activity must be evaluated using explicit computational hydrogen electrode (CHE) calculations. Considering the theoretical overpotential and the likelihood of observing a substantial fraction of active OER sites, MnCoNiCuZn is selected for further analysis in actual experimental catalytic environment.

Thermodynamic stability of MnCoNiCuZn

The thermodynamic stability for the proposed high entropy alloy catalyst MnCoNiCuZn, with stoichiometric ratio Mn:Co:Ni:Cu:Zn::2:2:8:12:12, is investigated. The atomic percentages of the elements Mn, Co, Ni, Cu, and Zn are 5.55%, 5.55%, 22.22%, 33.33%, and 33.33%, respectively. Investigations are carried out in a large bulk FCC crystal containing 108 atoms following the above-mentioned atomic percentages. Since HEAs are known to be stabilized due to their large mixing entropies,[39] 10 bulk HEA crystals with different randomized configurations are optimized and their mixing enthalpies/formation energies are computed to assess their thermodynamic stabilities in two different ways. The results are shown in Table I.

Here, the mixing enthalpy/formation energy, $\Delta H_{mix}^{relaxed}$, is obtained from bulk FCC HEA crystal, where both the lattice parameter and atomic coordinates are relaxed. In addition, ΔH_{mix}^{fcc} is obtained by fixing the lattice constant to that of an FCC crystal and only the atomic coordinates are optimized. In both cases, the thermodynamic criteria for formation of simple solid structures (-20 kJ mol^{-1} $\leq \Delta H_{mix} \leq 5$ kJ mol^{-1}) are satisfied.[23,40] A similar procedure is performed for the Gibbs free energy factor, $\Omega\left(=\frac{T_m \Delta S_{mix}}{|\Delta H_{mix}|}\right)$, where ΔH_{mix} is replaced. It is found that the computed values of the Gibbs free energy factor parameters, atomic radius fraction, and *VEC* for the proposed HEA composition are $\Omega^{relaxed}$, $\Omega^{fcc} \geq 1.1$, $\delta = 4.78\%$, and 10.77, respectively. The values consistent with the solid solution rules $\Omega \geq 1.1$ and $\delta \leq 6.6\%$, where *VEC* greater than 8 forms FCC configuration.[5] Therefore, the proposed HEA, MnCoNiCuZn, with the above-mentioned atomic percentage are expected to exist as stable FCC structures.

TABLE I. Thermodynamic viability of MnCoNiCuZn-HEA parameters for formation of the bulk FCC solid solution.

Configuration	$\Delta H_{\mathrm{mix}}^{relaxed}$ (kJ mol^{-1})	$\Delta H_{\mathrm{mix}}^{fcc}$ (kJ mol^{-1})	$\Omega^{relaxed}$	Ω^{fcc}
1	−4.81	−4.58	2.98	3.14
2	−4.95	−4.73	2.90	3.03
3	−5.23	−4.99	2.74	2.87
4	−4.84	−4.61	2.97	3.11
5	−4.54	−4.33	3.16	3.32
6	−4.35	−4.13	3.30	3.48
7	−4.67	−4.43	3.07	3.24
8	−5.27	−5.05	2.72	2.84
9	−4.86	−4.63	2.95	3.10
10	−5.16	−4.90	2.78	2.93

Simulating experimental environment

Catalyst surfaces in actual OER settings are known to reconstruct into (oxy)hydroxides, where MOOH (M = metal atom) acts as the functioning catalytic material.[7,41,42] Therefore, MnCoCuZnNiOOH is employed as the representative HEA catalyst model. Note that high-throughput screening is performed on FCC-based HEA crystals to identify promising elemental compositions and local motifs. However, the β-NiOOH model likely represents the reconstructed surface under operating conditions. Accordingly, the final OER activity is evaluated on an MnCoNiCuZn-doped β-NiOOH model. The catalyst is modeled to capture the interior of β − NiOOH, where the surface is doped with MnCoCuZn and the widely studied $(01\bar{1}2)$ facets are chosen for performing computational hydrogen electrode investigation. This pathway was considered with a practical picture in mind that more often than not, metallic HEAs are formed on top of nickel-foams that usually act as substrates.[11,43] To accurately capture the AEM steps for OER corresponding to experimental environment,[7,44] DFT calculations with Hubbard-U corrections were performed for this system. The previous PBEsol electronic exchange and correlation interactions along with the PAW method are maintained. Effective Hubbard-U parameters were taken from Wei *et al.*[44] that are optimized to match the $M(OH)_2$/MOOH redox features. The size of the chosen supercell is 3 × 2 × 3, where the bottom two layers were fixed. A vacuum of 15 Å was introduced in z−direction with 2 × 2 × 1 k−point grid.

For this particular system, the entire DFT+U calculation was performed for settings initialized at two different magnetic moments (μ_{int} = 2 and 4) as there is scatter in the atomic magnetic moment data.[45] Based on these settings, overpotentials were calculated in two ways: (Approach-1) In the first approach, the pristine surface is relaxed by initializing with two different magnetic moments. After relaxation, the most stable surface configuration is chosen for further computation. Using this slab, the energies of intermediates (*OH, *O, and *OOH) are evaluated to obtain the overpotential. (Approach-2) In the second approach, after initializing with two different magnetic moments, the lowest energy configuration for surface and surface + adsorbate (intermediates) systems are used to evaluate the overpotentials.

For oxygen evolution reactions through AEM pathways, the adjacent metal atoms M_1 and M_2 influence the reaction energies substantially. Therefore, the arrangement $M_1 - M_2{}^{\dagger} - M_3$ is considered, where the active site is denoted by the dagger symbol †. Recent studies suggest cobalt (Co) atoms as the dominant active site.[7] Therefore, focus is placed on M_2 = Co as the active site. The CHE model was carried out for ${}^5C_2 = 10$ different arrangements of $M_1 - M_2{}^{\dagger} - M_3$, where M_1, M_3 = Mn, Co, Ni, Cu, and Zn.

The CHE model consisting of the following four proton–electron transfer steps were considered:

$$^* + \mathrm{OH}^- \rightarrow {}^*\mathrm{OH} + \mathrm{e}^- \ (= \Delta G_1), \tag{5}$$

$$^{*}OH + OH^{-} \rightarrow {}^{*}O + H_2O(l) + e^{-} \ (= \Delta G_2), \quad (6)$$

$$^{*}O + OH^{-} \rightarrow {}^{*}OOH + e^{-} \ (= \Delta G_3), \quad (7)$$

$$^{*}OOH + OH^{-} \rightarrow {}^{*} + O_2 + H_2O(l) + e^{-} \ (= \Delta G_4). \quad (8)$$

The Gibbs free energy change (ΔG) of the intermediate was defined as[31]

$$\Delta G = \Delta E + \Delta ZPE + \Delta H_{vib} - T\Delta S. \quad (9)$$

The following formula was used to calculate the overpotential (η) for OER activity:

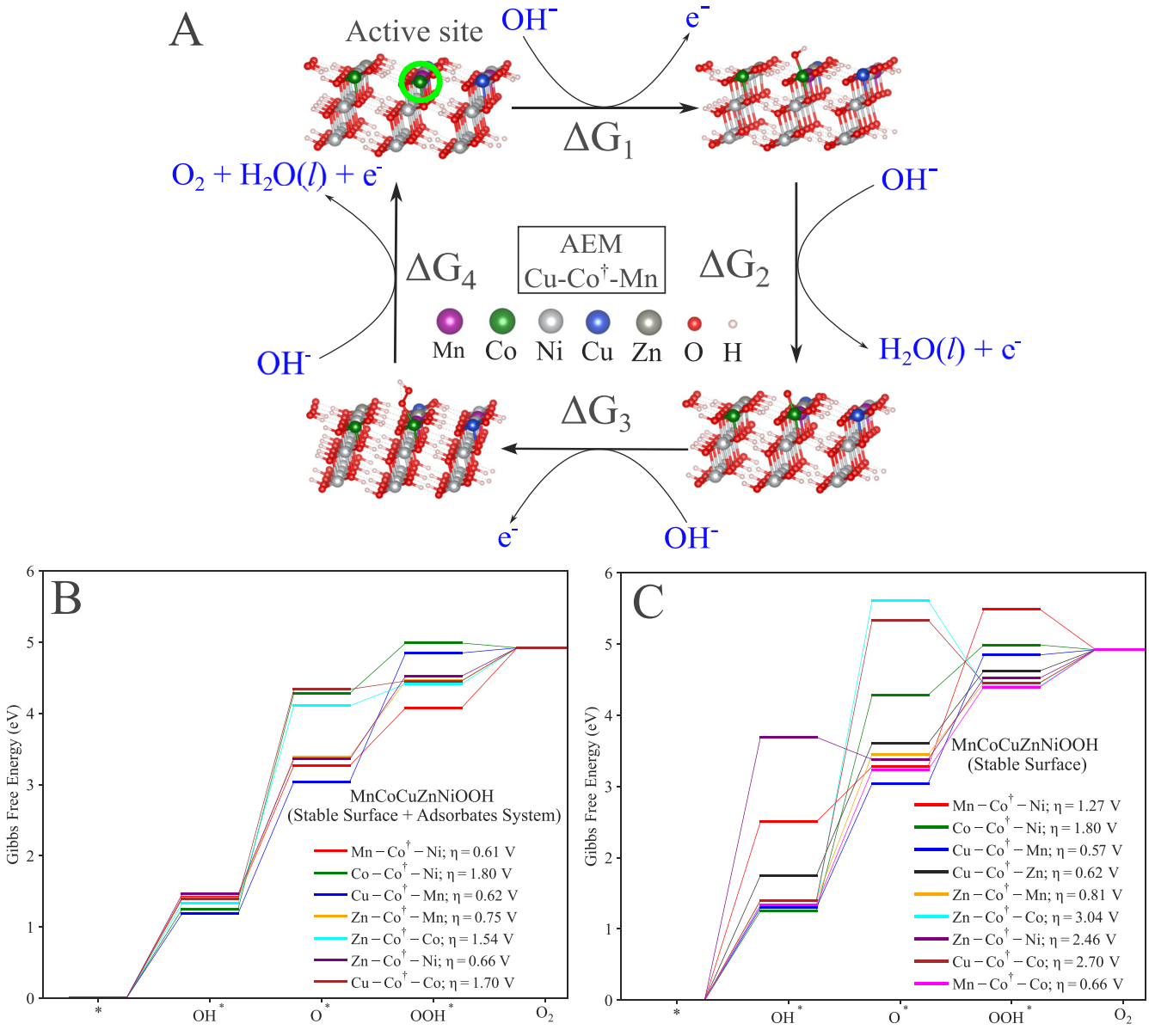

FIG. 4. AEM pathways and Gibbs free energy diagrams for OER, corresponding to the valid configurations (where reaction intermediates are stabilized) among the 10 different combinations of M_1-$Co^{\dagger}$-M_3 adsorption sites. The symbol † denotes the central Co atom as the active site, where M_1 and M_3 correspond to the neighboring metal atoms. (a) AEM pathway for oxygen evolution mechanism at Cu–$Co^{\dagger}$–Mn site discovered to exhibit lowest overpotential of (η =) 0.57 V corresponding to the most stable surface configuration. Gibbs free energy diagrams are obtained at 1.23 V vs RHE for two different scenarios: (b) Energies based on both stable surface + adsorbate configurations and (c) Based on energies of intermediate reactions occurring on top of the most stable surface configuration.

$$\eta = \frac{\max\left[\Delta G_1, \Delta G_2, \Delta G_3, \Delta G_4\right]}{e} - 1.23\ V. \tag{10}$$

As seen from Fig. 4, a substantial number of highly active sites were discovered that showed overpotential of nearly ~0.61 V. It was also found that when the Co site is adjacent to Cu and Mo, forming a Cu–Co†–Mn site, the overpotential can be as low as 0.57 V. Note that the reported overpotentials are 0.61–0.67 V[37,38] for IrO_2 in the literature. However, using identical computational settings to our work, the obtained overpotential was 0.33 V for IrO_2 (settings discussed in the supplementary material), indicating that IrO_2 still exhibits better performance than the proposed catalyst.

In addition, verification of the scaling relationship[46] between OH* and OOH* adsorption free energies for the specific site of Cu–Co†–Mn, where overpotential resulted in 0.57 V, was also carried out. Note that from actual computation, $\Delta G_{OOH,DFT}$ = 4.85 eV and $\Delta G_{OH,DFT}$ = 1.29 eV, while from scaling relationship, $\Delta G_{OOH,scaling}$ = 4.49 eV. This leads to the following relations: $\Delta G_{OOH,DFT} = \Delta G_{OH,DFT} + 3.56$, which indicates deviations from the scaling relationship.

As the proposed HEA contains Zn and Cu, it is reasonable to suspect susceptibility to dissolution under anodic alkaline environments. To evaluate the stability, Pourbaix diagram is constructed, using equilibrium aqueous species data consistent with what is reported in Materials Project. The proposed catalyst, $Mn_2Co_2Ni_8Cu_{12}Zn_{12}$, in the β – NiOOH phase lies in a stable region at 0.97 and 1.01 V vs SHE (for η = 0.57 and 0.61 V and vs RHE) at pH 14, which indicates stability. The details of the Pourbaix diagram, along with the reactions can be found in the supplementary material.

CONCLUSION

Machine learning assisted *ab initio* computational DFT/DFT+U screening of HEA catalysts was performed. A roughly 23 000 OH-adsorption energy dataset was developed for FCC (111) crystal surfaces comprising of 5 elements chosen from Mn, Fe, Co, Ni, Cu, Zn, and Mo. Generative machine learning models such as autoencoder architecture was found to capture the clustering of HEA configurations based on their adsorption energies. However, the mean absolute errors for test and train data were comparable with a simple linear regression model. The linear regression model was iteratively trained for six cycles and employed to screen more than 163×10^9 configuration space comprising the entire HEA types. It was shown that using a theoretical overpotential derived from OH binding energy computed from DFT and linear scaling relationship promising OER catalyst was found to be most likely from MnCoNiCuZn-HEA type. This is also consistent with the two best candidates that resulted from screening 1×10^6 configuration space of 21 HEAs with 8725 stoichiometry each. It was found that although theoretical overpotential calculation resulted in an overpotential of 0.67 V, deviations from explicit computation revealed that it could be as high as 0.99 V. Later, to emulate actual experimental catalytic environment, HEA elemental composition was doped on top of the widely used *beta*–NiOOH ($01\bar{1}2$) facet. Overpotential resulting from explicit CHE simulations gives η of 0.61 V. However, when the Co site becomes adjacent to Cu and Mn, forming a Cu–Co†–Mn site, the overpotential becomes as low as 0.57 V. Analysis indicates that upon consistent comparison, the most stable rutile IrO_2 (η = 0.33 V) exhibits superior performance relative to the proposed MnCoNiCuZn catalyst (η = 0.57 V). However, the present methodology, which resulted in the proposed catalyst owing to its earth-abundant composition and the presence of statistically favorable adsorption sites, could successfully identify high probability density of discovering highly active sites within a compositional space.

SUPPLEMENTARY MATERIAL

The supplementary material provides details on various machine learning models, sampling techniques, and other modeling details employed in this article.

ACKNOWLEDGMENTS

The authors acknowledge the support from Amazon–IIT Bombay AI/ML Initiative (AIDo-AIAIMLI) under research Grant No. 00000000000010019448 (Program RD/0125-AMAZO09-002). The authors acknowledge the additional financial support from the IRCC, IIT Bombay, and the National Supercomputing Mission, Government of India. Computational resources were provided by the SpaceTime-II, SpaceTime-III, and PARAM RUDRA supercomputing facilities at IIT Bombay.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Jyotishraj Thoudam: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Ankit Jain**: Conceptualization (equal); Data curation (equal); Funding acquisition (equal); Investigation (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] B. Cantor, I. T. H. Chang, P. Knight, and A. J. B. Vincent, "Microstructural development in equiatomic multicomponent alloys," Mater. Sci. Eng. A **375–377**, 213–218 (2004).
[2] J.-W. Yeh, S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, and S.-Y. Chang, "Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes," Adv. Eng. Mater. **6**, 299–303 (2004).
[3] E. P. George, D. Raabe, and R. O. Ritchie, "High-entropy alloys," Nat. Rev. Mater. **4**, 515–534 (2019).
[4] C. Oses, C. Toher, and S. Curtarolo, "High-entropy ceramics," Nat. Rev. Mater. **5**, 295–309 (2020).

[5]X. Yang and Y. Zhang, "Prediction of high-entropy stabilized solid-solution in multi-component alloys," Mater. Chem. Phys. **132**, 233–238 (2012).
[6]T. A. A. Batchelor, J. K. Pedersen, S. H. Winther, I. E. Castelli, K. W. Jacobsen, and J. Rossmeisl, "High-entropy alloys as a discovery platform for electrocatalysis," Joule **3**, 834–845 (2019).
[7]Y. Mei, J. Chen, Q. Wang, Y. Guo, H. Liu, W. Shi, C. Lin, Y. Yuan, Y. Wang, B. Y. Xia, and Y. Yao, "MoZn-based high entropy alloy catalysts enabled dual activation and stabilization in alkaline oxygen evolution," Sci. Adv. **10**, eadq6758 (2024).
[8]O. Vidal, B. Goffé, and N. Arndt, "Metals for a low-carbon society," Nat. Geosci. **6**, 894–896 (2013).
[9]E. Lèbre, M. Stringer, K. Svobodova, J. R. Owen, D. Kemp, C. Côte, A. Arratia-Solar, and R. K. Valenta, "The social and environmental complexities of extracting energy transition metals," Nat. Commun. **11**, 4823 (2020).
[10]S. H. Ali *et al.*, "Mineral supply for sustainable development requires resource governance," Nature **543**, 367–372 (2017).
[11]X. Yan, Y. Zhou, and S. Wang, "Nano-high entropy materials in electrocatalysis," Adv. Funct. Mater. **35**, 2413115 (2025).
[12]Q. Wang and Y. Yao, "Harnessing machine learning for high-entropy alloy catalysis: A focus on adsorption energy prediction," npj Comput. Mater. **11**, 91 (2025).
[13]Y. Yao, Q. Dong, A. Brozena, J. Luo, J. Miao, M. Chi, C. Wang, I. G. Kevrekidis, Z. J. Ren, J. Greeley, G. Wang, A. Anapolsky, and L. Hu, "High-entropy nanoparticles: Synthesis-structure-property relationships and data-driven discovery," Science **376**, eabn3103 (2022).
[14]I. C. Man, H. Su, F. Calle-Vallejo, H. A. Hansen, J. I. Martínez, N. G. Inoglu, J. Kitchin, T. F. Jaramillo, J. K. Nørskov, and J. Rossmeisl, "Universality in oxygen evolution electrocatalysis on oxide surfaces," ChemCatChem **3**, 1159–1165 (2011).
[15]J. Rossmeisl, A. Logadottir, and J. K. Nørskov, "Electrolysis of water on (oxidized) metal surfaces," Chem. Phys. **319**, 178–184 (2005).
[16]J. Rossmeisl, Z.-W. Qu, H. Zhu, G.-J. Kroes, and J. K. Nørskov, "Electrolysis of water on oxide surfaces," J. Electroanal. Chem. **607**, 83–89 (2007).
[17]G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science **313**, 504–507 (2006).
[18]G. Kresse and J. Furthmüller, "Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set," Phys. Rev. B **54**, 11169 (1996).
[19]J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard, and H. Jónsson, "Origin of the overpotential for oxygen reduction at a fuel-cell cathode," J. Phys. Chem. B **108**, 17886–17892 (2004).
[20]C. Kittel, *Introduction to Solid State Physics*, 8th ed. (John Wiley & Sons, New York, 2004).
[21]N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976).
[22]A. H. Larsen *et al.*, "The atomic simulation environment—a Python library for working with atoms," J. Phys.: Condens. Matter **29**, 273002 (2017).
[23]S. Guo, "Phase selection rules for cast high entropy alloys: An overview," Mater. Sci. Technol. **31**, 1223–1230 (2015).
[24]L. Jiang, Y. P. Lu, H. Jiang, T. M. Wang, B. N. Wei, Z. Q. Cao, and T. J. Li, "Formation rules of single phase solid solution in high entropy alloys," Mater. Sci. Technol. **32**, 588–592 (2016).
[25]F. d. Boer, W. Mattens, R. Boom, A. Miedema, and A. Niessen, *Cohesion in Metals* (North-Holland, 1988), Vol. 1.
[26]A. Takeuchi and A. Inoue, "Classification of bulk metallic glasses by atomic size difference, heat of mixing and period of constituent elements and its application to characterization of the main alloying element," Mater. Trans. **46**, 2817–2829 (2005).
[27]S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu," J. Chem. Phys. **132**, 154104 (2010).
[28]G. Kresse and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method," Phys. Rev. B **59**, 1758–1775 (1999).
[29]J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, "Restoring the density-gradient expansion for exchange in solids and surfaces," Phys. Rev. Lett. **100**, 136406 (2008).
[30]S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, "Electron-energy-loss spectra and the structural stability of nickel oxide:an LSDA+U study," Phys. Rev. B **57**, 1505–1509 (1998).
[31]B. Zhang *et al.*, "Homogeneously dispersed multimetal oxygen-evolving catalysts," Science **352**, 333–337 (2016).
[32]D. Visaria and A. Jain, "Machine-learning-assisted space-transformation accelerates discovery of high thermal conductivity alloys," Appl. Phys. Lett. **117**, 202107 (2020).
[33]I. Batatia, S. Batzner, D. P. Kovács, A. Musaelian, G. N. C. Simm, R. Drautz, C. Ortner, B. Kozinsky, and G. Csányi, "The design space of E(3)-equivariant atom-centred interatomic potentials," Nat. Mach. Intell. **7**, 56–67 (2025).
[34]A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, "Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons," Phys. Rev. Lett. **104**, 136403 (2010).
[35]A. P. Bartók, R. Kondor, and G. Csányi, "On representing chemical environments," Phys. Rev. B **87**, 184115 (2013).
[36]M. Sokolov and K. S. Exner, "Is the $*$O vs. $*$OH scaling relation intercept more relevant than the $*$OOH vs. $*$OH intercept to capture trends in the oxygen evolution reaction?," Chem Catal. **4**, 101039 (2024).
[37]J. Feng, Z. Dong, Y. Ji, and Y. Li, "Accelerating the discovery of metastable IrO_2 for the oxygen evolution reaction by the self-learning-input graph neural network," JACS Au **3**, 1131–1140 (2023).
[38]Z. Fan, Y. Ji, Q. Shao, S. Geng, W. Zhu, Y. Liu, F. Liao, Z. Hu, Y.-C. Chang, C.-W. Pao, Y. Li, Z. Kang, and M. Shao, "Extraordinary acidic oxygen evolution on new phase 3R-iridium oxide," Joule **5**, 3221–3234 (2021).
[39]B. C. Wyatt *et al.*, "Order-to-disorder transition due to entropy in layered and 2D carbides," Science **389**, 1054–1058 (2025).
[40]D. Alloyeau *et al.*, "Synthesizing high-entropy alloy nanoparticles: General discussion," Faraday Discuss. **264**, 178 (2026).
[41]Z.-F. Huang, J. Song, Y. Du, S. Xi, S. Dou, J. M. V. Nsanzimana, C. Wang, Z. J. Xu, and X. Wang, "Chemical and structural origin of lattice oxygen oxidation in Co–Zn oxyhydroxide oxygen evolution electrocatalysts," Nat. Energy **4**, 329–338 (2019).
[42]E. Fabbri, M. Nachtegaal, T. Binninger, X. Cheng, B.-J. Kim, J. Durst, F. Bozza, T. Graule, R. Schäublin, L. Wiles, M. Pertoso, N. Danilovic, K. E. Ayers, and T. J. Schmidt, "Dynamic surface self-reconstruction is the key of highly active perovskite nano-electrocatalysts for water splitting," Nat. Mater. **16**, 925–931 (2017).
[43]W. Zheng, M. Liu, and L. Y. S. Lee, "Best practices in using foam-type electrodes for electrocatalytic performance benchmark," ACS Energy Lett. **5**, 3260–3264 (2020).
[44]L. Wei, M. D. Hossain, M. J. Boyd, J. Aviles-Acosta, M. E. Kreider, A. C. Nielander, M. B. Stevens, T. F. Jaramillo, M. Bajdich, and C. Hahn, "Insights into active sites and mechanisms of benzyl alcohol oxidation on nickel–iron oxyhydroxide electrodes," ACS Catal. **13**, 4272–4282 (2023).
[45]M. K. Horton, J. H. Montoya, M. Liu, and K. A. Persson, "High-throughput prediction of the ground-state collinear magnetic order of inorganic materials using density functional theory," npj Comput. Mater. **5**, 64 (2019).
[46]A.-I. Hutu, E. Pervolarakis, I. N. Remediakis, H. H. Kristoffersen, and J. Rossmeisl, "Scaling relations on high-entropy alloy catalyst surfaces," J. Phys. Chem. C **128**, 10251–10258 (2024).