Exploring the Relationship between the CO Oxidation Reaction and the Elemental Composition in Electrocatalysts with Machine Learning - a Case Study of PtRuPdRhAu High Entropy Alloys

Vladislav A. Mints,[†] Jack K. Pedersen,[‡] Alexander Bagger,[‡] Jonathan Quinson,[‡] Andy S. Anker,[‡] Kirsten M. Ø. Jensen,[‡] Jan Rossmeisl,[‡] and Matthias Arenz[*,†]

†*Department of Chemistry, Biochemistry and Pharmaceutical Sciences, University of Bern, Bern, Switzerland*

‡*Department of Chemistry, University of Copenhagen, Copenhagen, Denmark*

E-mail: matthias.arenz@unibe.ch

Abstract

In recent years, the development of complex multi-metallic nanomaterials such as high entropy alloy (HEA) catalysts has gained popularity. Composed of 5 or more metals, the compositions of HEAs exhibit extreme diversity. Their diversity provides both a promising avenue to identify new catalysts and is at the same time a severe constraint on their preparation and study. To address the challenges related to the preparation, study and optimization of HEAs, machine learning solutions are attractive. In this paper, the composition of PtRuPdRhAu hydrogen oxidation catalysts is optimized for the CO oxidation reaction.

This is achieved by constructing a dataset using Bayesian optimization as guidance. For this quinary nanomaterial, the best performing composition was found within the first 35 experiments. However, the dataset was expanded until a total of 68 samples were investigated. This final dataset was used to construct a random forest regression model and a linear model. These machine learned models were used to assess the relationships between the concentrations of the constituent elements and the CO oxidation reaction onset potential. The onset potentials were found to correlate with the composition dependent adsorption energy of *OH obtained from density functional theory. This study demonstrates how machine learning can be employed in an experimental setting to investigate the vast compositional space of HEAs.

Introduction

Achieving a carbon neutral society through fossil fuel free technologies is one of the main challenges faced by mankind today. Within the scope of this challenge, the proton exchange membrane fuel cell has been developed.[1–3] In this specific fuel cell, carefully selected catalysts carry out the hydrogen oxidation reaction (HOR) and the oxygen reduction reaction (ORR) to generate electricity. The ORR is a major topic for studies as it requires large overpotentials to take place.[4] On the other hand, the hydrogen reaction is efficiently catalyzed by platinum with neglectable overpotentials.[5] However, platinum is highly susceptible to CO poisoning, whose traces are present in the hydrogen gas from the widely employed water gas shift reaction or potentially from on board methanol reformers.[6–8] Therefore, it is still an ongoing quest to find a HOR catalyst that shows resilience towards CO poisoning.[9]

A common strategy to increase the CO tolerance of platinum is to alloy it with ruthenium.[10] This produces a bifunctional catalyst that can perform both the HOR and the CO oxidation at low overpotentials.[11] The proposed mechanism for PtRu mediated catalysis is that ruthenium facilitates the formation of adsorbed OH species at lower potentials.[12] These adsorbed OH species can oxidize the CO adsorbed on platinum which leads to the reactiva-

tion of the surface. In addition, it has been proposed that an electronic effect of ruthenium weakens the CO bond on platinum.[13]

Studies on alloy catalysts are conventionally limited to bi- and trimetallic alloys. However, in very recent years the topic of high entropy alloys (HEAs) has become popular.[14–16] HEAs contain five or more elements in a randomized order, which produces a surface with multiple different active sites.[17] Further, the diversity of elements with different sizes distorts the crystal lattice which gives HEAs not only unique mechanical properties but also changes the electronic structure.[18] This change in electronic structure is expected to provide unique catalytic properties that have yet to be discovered.[19]

The experimental and theoretical investigations of HEAs form a major challenge due to the immense amount of available compositions. In an HEA containing five elements there are $\frac{99!}{95!4!} = 3,764,376$ compositions possible when each elemental concentration is varied between 1 at.% (atomic percent) and 96 at.% only by 1 at.%. This amount of possible compositions cannot be investigated experimentally in a timely manner. Consequently, new strategies to approach the study and discovery of relevant HEAs with dedicated properties are necessary.[20,21] Currently, one approach is focused on constructing a methodology employing density functional theory (DFT) that enables prediction of highly active HEA compositions, which then are evaluated experimentally.[22,23] Another approach, which is employed in this paper, is based on combining screening experiments with machine learning.[24,25]

Machine learning is slowly spreading into fuel cell and electrocatalysis reasearch.[26] One application in fuel cell research is prediciting the stability of fuel cells.[27,28] However, machine learning has also been employed to optimize the synthesis protocol to construct Ni-rich cathode materials for Li-ion batteries.[29] Lately, machine learning has been used to optimize the PtPdAu alloy for methanol oxidation.[24] In future years, the use of machine learning in nanomaterials discovery is only expected to rise.[30]

In this research, PtRuPdRhAu has been studied as a CO resistant HOR catalyst. Pt and Ru were included in this HEA, since they form a well known bifunctional catalyst.[31] Au

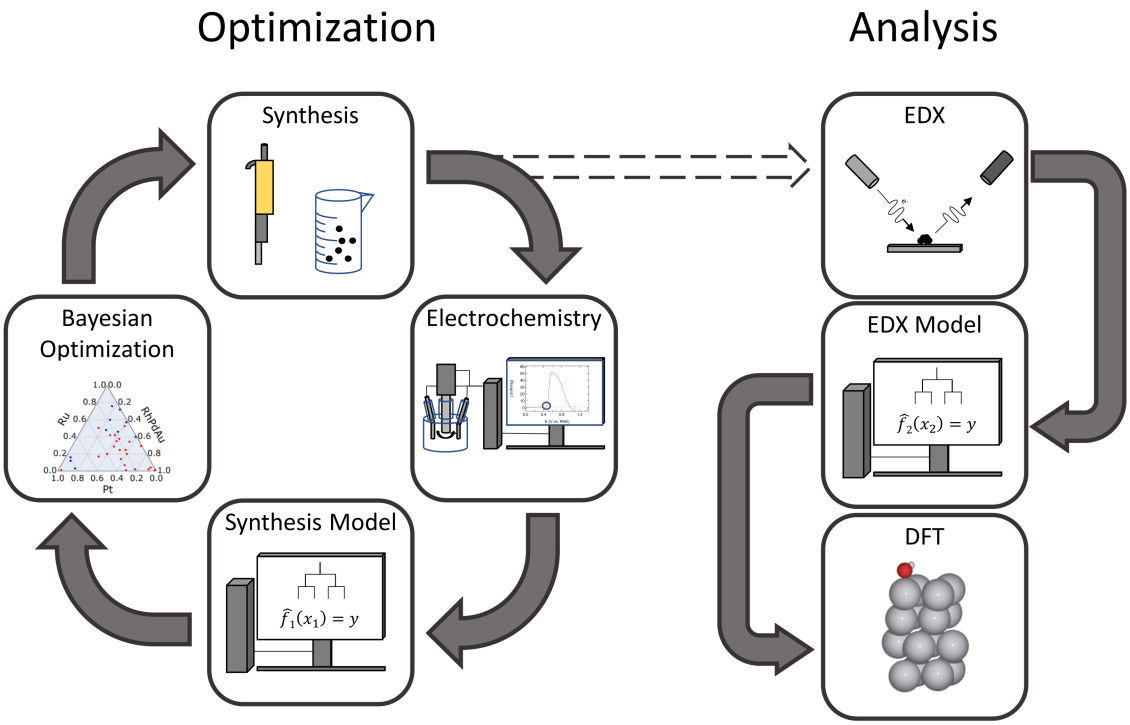

Figure 1: Schematic representation of the optimization loop and the analysis steps.

was added as it is a known CO oxidation catalyst in alkaline media.[32] Furthermore, Au was used to facilitate the synthesis of the nanoparticles. Rh and Pd by themselves show weak activities towards the CO oxidation.[33,34] However, in combination with Pt, this trio forms the three-way catalyst which is widely used to remove CO from exhaust gasses.[35,36]

The conceptual design of this study of PtRuPdRhAu, which was carried out in a screening setting, is summarized in Figure 1. In the first part, the composition was optimized for the CO oxidation reaction. This search was carried out using Bayesian optimization, similar to the approach of Nugraha et al.[24] The ratio of the precursor mixtures was used as the input parameter for the optimization algorithm. The affinity of the alloy towards CO oxidation was used as the output parameter. In the second part, machine learned models were constructed with a new input parameter, namely the compositions that were obtained from energy dispersive X-ray spectroscopy (EDX). These models were used to elucidate the relationship between the concentration of the elements and the CO oxidation reaction. Finally,

the correlations were related to physical phenomena using DFT models.

Results and Discussion

Synthesis

In previous research a surfactant-free synthesis method in alkaline mono-alcohols of colloidal precious metal nanomaterials has been developed.[37,38] This synthesis is suitable to synthesize catalytically active mono- or bimetallic nanoparticles at room temperature.[39] It was observed that at room temperature, and without surfactant, the same reduction occurs in the presence of Au, as it could be expected from previous reports.[40] The presence of a support limited the formation of macroscopic Au particles, yet relatively large Au nanoparticles were formed ($<$ 20 nm) when only a Au precursor like $HAuCl_4$ was used. However, in the presence of other complexes comprising elements like Pt, Ir, Ru, Rh, and Pd, a reaction proceeds very fast at room temperature, as shown in Figure S1. The reduction of the Au precursor in the presence of the additional precursors leads to the formation of small nanoparticles, ca. 3 nm in diameter, as shown in Figure S2. The formed nanoparticles are only slightly larger than monometallic Pt, Ir, or Pd nanoparticles obtained in alkaline methanol which is an encouraging sign towards the formation of multi-metallic nanomaterials.[37,41]

In the present work, we applied this surfactant-free room temperature synthesis for the first time as a fast and simple production method of carbon supported multi-metallic nanomaterials. The chloride based precursors of the used metals were mixed together with carbon flakes in alkaline methanol and left stirring for 10 minutes. This resulted in the production of highly dispersed supported nanoparticles. The formation of nanoparticles was proved in transmission electron microscopy (TEM) for the best performing catalyst. Its TEM micrographs can be observed in Figure 2. With the help of EDX, the incorporation of Pt, Ru, Pd, Rh, and Au in the nanostructures was confirmed.

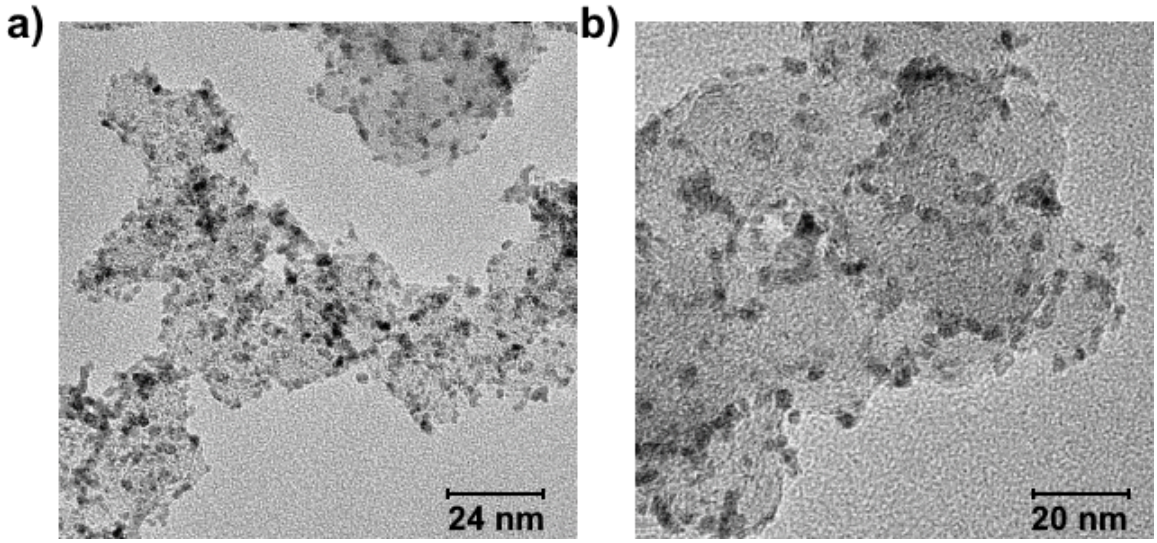

Figure 2: **a-b)** TEM micrographs of the sample $Pt_{12}Ru_{38}Pd_8Rh_{41}Au_1$.

Bayesian optimization

The samples for investigation were selected using Bayesian optimization. Bayesian optimization is a tool that allows a guided search through an unkown multidimensional space.[24,25] Therefore, it can be employed in combination with experiments to quickly identify a catalyst with an optimal activity. In this research, this optimization is used to find the most active composition within the PtRuPdRhAu space for the CO oxidation reaction. The Bayesian optimization was performed in conjunction with a random forest regression model that was trained on the experimental data to provide multiple suggestions for further investigation. In order to construct this regression, it is important to first consider which parameters are used as the input and output parameters.

Since the scope of this research is to find the best performing catalyst, the output parameter of the model must be a quantitative expression for the CO oxidation affinity. This measure for the CO oxidation affinity will be further addressed with the CO oxidation onset potential. The CO oxidation onset potential was measured electrochemically by depositing a monolayer of CO on top of the nanoparticle, which was subsequently oxidized by an anodic sweep in an H_2 rich environment. The HOR functioned as a probe reaction to amplify the

potential at which the CO monolayer was oxidized. We defined the CO oxidation as the point where an increase of 1.5 mA $mg^{-1}_{PtRuPhRhAu}$ on top of the capactive current was measured. As the CO oxidation was evaluated with the HOR as a probe reaction, only catalysts that were active for the HOR above the CO oxidation onset potential were taken into account in the machine learning models.

The input parameters for the model should reflect the properties of the HEA that influence the CO oxidation activity. For this, the most straightforward choice is the elemental composition of the HEA. However, the final composition of the prepared nanomaterials cannot be systematically and timely assessed. Arguably, the most controlled parameters that characterize a given PtRuPdRhAu particle is the amount of each precious metal used during the synthesis. Therefore, the at.% of metals used in the synthesis was used as an input parameter. All models that used the synthesis precursor ratio as input will be denoted with *synthesis models*.

The inital 25 samples were randomly selected and fed to the Bayesian optimization algorithm. This algorithm suggested 10 new compositions, which were experimentally investigated and added to the dataset. This addition of 10 new samples concludes a single optimization cycle. In total five optimization cycles were carried out. The summary of this Baysian optimization process is shown in Figure 3. Figure 3a-c shows a representation of the sampling of each Bayesian optimization cycle. It can be observed that after the second cycle of optimization, the computational methods start to select samples with large Ru content. Figure 3d shows the average CO oxidation onset potentials for the investigated samples. It concludes that the most active catalyst composition, with a precursor ratio of $Pt_{16}Ru_{46}Pd_2Rh_{34}Au_2$, was found within the first optimization cycle at the 32nd entry. The composition of this alloy, obtained with EDX, was $Pt_{12}Ru_{38}Pd_8Rh_{41}Au_1$. It is also observed that in the subsequent 2rd, 3th and 4th cycles the onset potentials were around the maximum of 0.47 V vs. RHE. As the samples in the 3rd-5th cycle have on average large Ru contents, it can be implied that Ru is playing a major role in the CO oxidation reaction. In the 5th

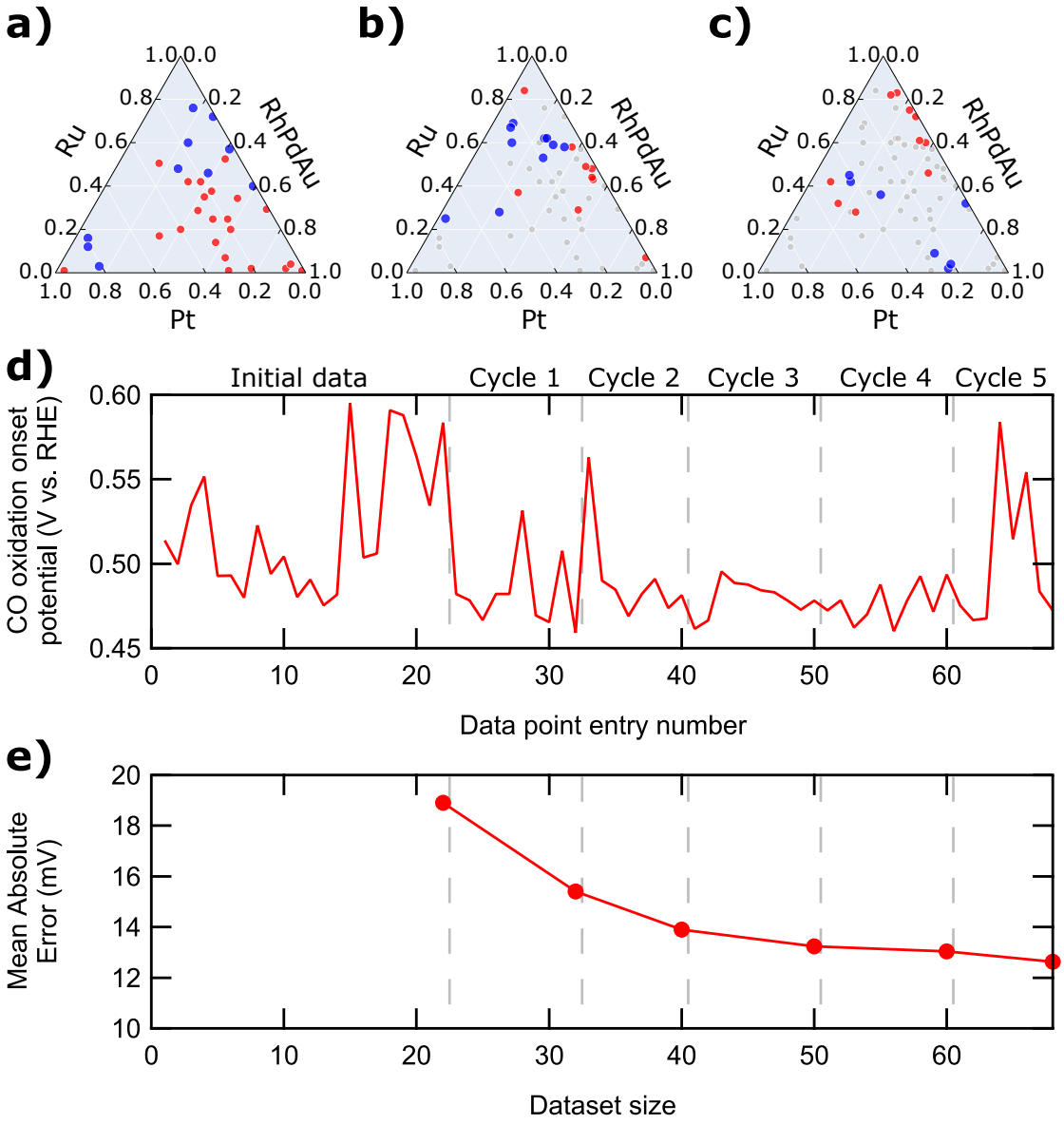

Figure 3: **a-c)** Representation of the compositions experimentally investigated in each Bayesian optimization cycle, with Rh-Pd-Au contracted to a single concentration. **a)** cycle 1 (red), cycle 2 (blue); **b)** cycle 3 (red), cycle 4 (blue), cycle 1-2(grey); **c)** cycle 5 (red), cycle 6 (blue), cycle 1-4 (grey). **d)** The experimentally determined CO oxidation onset potential for each data point in the investigated order. The dashed line indicate the regions of each Bayesian optimization cycle. **e)** The change in the mean absolute error, obtained from the leave-one-out cross-validation method, of the synthesis random forest regression model with each Bayesian optimization cycle.

cycle, several low performing compositions were selected for investigation, to explore further poorly represented regions. This exploration produced samples with low CO tolerance, in line with the prediction. This proposes that the optimization has already found the global maximum in the first cycle. Finally, the evolution of the synthesis random forest regression models can be be observed in Figure 3e. This evolution is represented by the mean absolute error (MAE) which was calculated with leave-one-out cross-validation (LOOCV). This MAE showed an exponential decrease with increasing dataset size. The final synthesis model had a MAE of 13 mV.

Machine learning

While the precursor ratio serves as an excellent input parameter for the Bayesian optimization procedure, it is poorly reflecting the true catalyst compositions. As such, the obtained correlations between the precursor ratios and the CO oxidation onset might not be as informative as required to understand in more depth this relationship. Therefore, all sampled catalysts were evaluated with EDX, the results of which were used to construct new machine learned models. These models, that use the EDX evaluated composition as input value will be referred to as *EDX models*.

With this EDX corrected dataset, two intrinsically different machine learned models were created. The first model was the EDX random forest regression, which is highly flexible in its nature as it has no bias on the expected correlations. The second is a more rigid EDX linear model. The linear model was limited to the first and second degree polynomial terms to avoid overfitting with higher degree terms. The obtained coefficients were corrected to account for the statistical probability of finding a specific elemental arrangement at the surface. This model is outlined in Equation 1. In this equation $E(X)$ is the CO oxidation onset potential, f is the fraction of an element m in at.%, β is the fitted and corrected coefficient, $P(X)$ is the probability of finding a specific surface arangement. Nevertheless, this model has a total of 21 predictors, which would severely overfit the small dataset of 68 points by using the

ordinary least squares regression. Therefore, in order to reduce the amount of predictors, the model was fitted using the lasso regression.[42]

$$
\begin{aligned}
E(f_{\mathrm{Pt}}, f_{\mathrm{Ru}}, f_{\mathrm{Pd}}, f_{\mathrm{Rh}}, f_{\mathrm{Au}}) &= \beta_0 + \sum_i \beta_{m_i} P(f_{m_i}) + \sum_{i \geq j} \beta_{m_i m_j} P(f_{m_i} f_{m_j}) \\
P(f_{m_i}) &= 0.5 f_{m_i} \\
P(f_{m_i} f_{m_j}) &= \begin{cases} f_{m_i} f_{m_j} & \text{if } i > j \\ 0.5 f_{m_i} f_{m_j} & \text{if } i = j \end{cases}
\end{aligned}
\tag{1}
$$

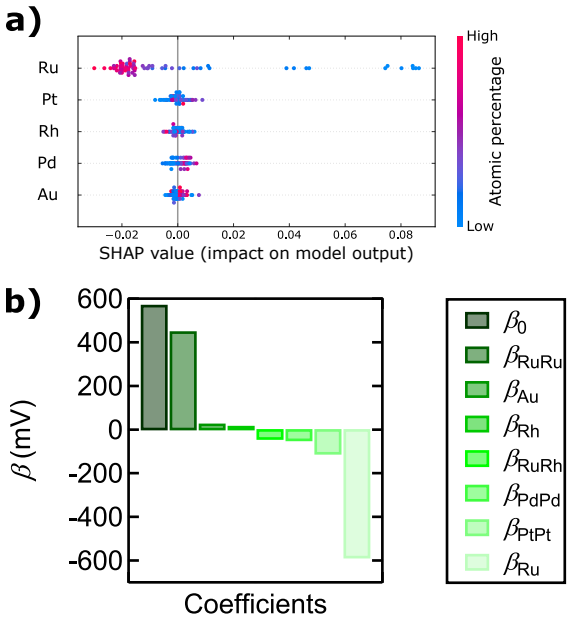

Figure 4: **a)** SHAP analysis of the EDX random forest regression model. **b)** The coefficients of EDX Lasso regression model.

The interpretation of the EDX random forest model, was achieved using SHapely Additive exPlanation (SHAP), the results of which are shown in Figure 4a.[43,44] The SHAP value represents the perturbation of the CO oxidation onset potential by the element from the average onset potential. Hence negative SHAP values are associated with a decrease in overpotential for the CO oxidation. It turns out that Ru has the largest impact on the CO oxidation. In low quantities Ru exhibits a very large positive SHAP value and at large

quantities a negative SHAP value. This is congruent with the already reported activity of Ru towards CO oxidation being much higher compared to the other investigated metals.[45] The next elements that show to improve the CO oxidation in large quantities are Pt and Rh. These two metals have also been shown to oxidize the CO monolayer at higher potentials compared to Ru, however at lower potentials than Pd.[34] In contrast to Ru, Pt, and Rh, large quantities of Pd showed to increase the CO oxidation onset potential, relative to the average, in large quantities as shown in Figure 4a. Au, in comparison to the other elements, did not show to have a large impact on the CO oxidation. This could be explained by a different mechanism for the formation of the CO monolayer on Au. While CO is forming a stable chemisorbed monolayer on Pt, Pd, Rh, and Ru, CO does not form a stable monolayer on gold in a CO deficient environment.[32] In addition, Au is inactive for the HOR.[46] Therefore, Au might be inactive for the CO oxidation in the current experimental setting.

The terms of the lasso regression with their coefficients can be observed in Figure 4b. Similar to the random forest model, the lasso regression shows that the term $P(f_{\mathrm{Ru}})$ has the largest associated negative coefficient. Therefore, Ru has the biggest impact on reducing the onset potential. In addition, linear models show that $P(f_{\mathrm{Ru}}f_{\mathrm{Ru}})$ has a large associated positive coefficient. This suggests that there is an optimum amount of Ru in an alloy. The next term that improves the CO oxidation is $P(f_{\mathrm{Pt}}f_{\mathrm{Pt}})$, which is also in agreement with the random forest model. When the linear model is used to predict the performance for a PtRu alloy, an optimum is found around the $\mathrm{Ru_{52}Pt_{48}}$ composition. This composition overlaps with previous studies that observed a similar maximum for this binary alloy.[31,47] Finally, the lasso regression includes the terms: $P(f_{\mathrm{Rh}})$, $P(f_{\mathrm{Au}})$, $P(f_{\mathrm{Ru}}f_{\mathrm{Rh}})$, $P(f_{\mathrm{Pd}}f_{\mathrm{Pd}})$ in the fit. However, their coefficients are too small to draw a solid conclusion upon.

The analysis of the SHAP values calculated on the EDX model concluded that among the investigated five elements Ru, Pt and Rh are lowering the CO oxidation onset potential the most. This allows a visualization of the activity of this hypothetical RuPtRh ternary alloy. In Figure 5a and b, the CO oxidation onset potential of the RuPtRh alloys are predicted

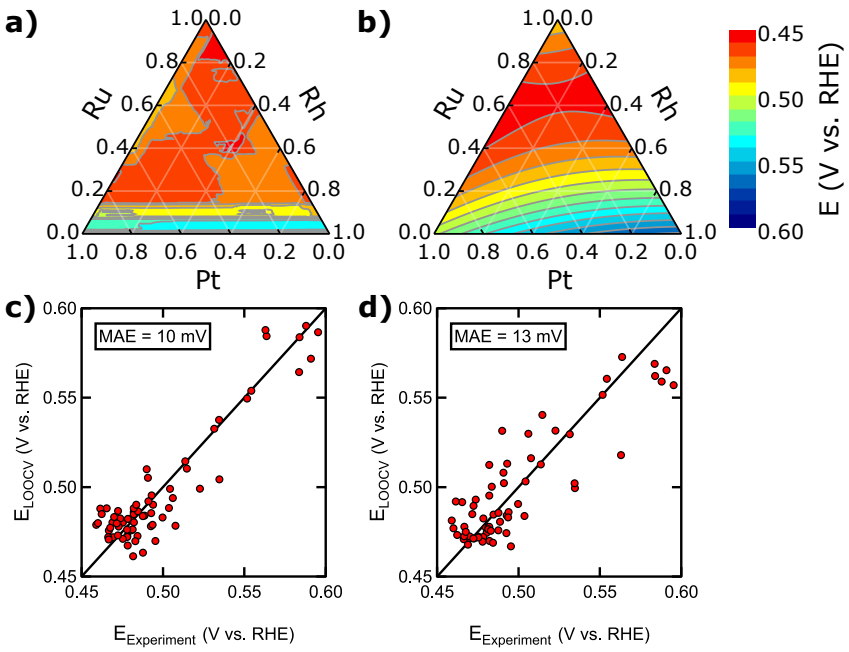

Figure 5: **a-b)** The slice of the hyperspace spanned by Ru, Pt and Rh and **c-d)** the prediction of the LOOCV vs the experimental value of the models: **a,c)** EDX Random Forest Regression **b,d)** EDX Lasso Regression.

using the EDX random forest regression and EDX lasso regression respectively. Both EDX models show that there is not a single most active alloy, but different regions which contain alloys that exhibit similar high activities. The shapes of these regions differ between the two EDX models due to their intrinsic flexibility. The lasso regression assumes a quadratic relation ship, which produces straight and smooth boundaries for the maximum domain. In comparison, the random forest regression has no prior bias, which produces domains with an irregular boundary.

Finally, the performance of the two EDX models was compared using the MAE of the LOOCV. This evaluation is shown in Figure 5c-d. The EDX random forest regression model outperforms the lasso regression slightly with an MAE of 10 mV. This is highly correlated with the high flexibility of the model. On the other hand the lasso regression has an MAE of 13 mV. It should be noted that the investigated compositions were selected using an optimization algorithm, which focused mainly on samples in the high performance region.

As observed, this region consists mostly of alloys with large concentrations of Ru. Therefore, the models in this study are capable of predicting the activity of catalysts that fall into the PtRuPdRhAu space with large Ru quantities. However, the error increases for alloys with little to no Ru content.

Simulations with density functional theory (DFT)

In order to gain a pysical understanding of the dependence of the onset potential on the composition, DFT simulations were conducted. The task of finding the potential limiting step of the reaction enabling HOR is formidable given that the surface structures in this work are not well defined. Additionally, the many possible adsorption sites on the HEA surface contribute to an inherent complexity in the variation of active sites. This further hinders a successful and timely discovery of the responsible chemical step. Therefore, we pursued a predictive model which could explain the trend in the observed onset potentials through electronic structure insights. This happened at the cost of loosing the explanation for the onset at the atomic level.

Figure 6 shows the overall reasonable predictive trend obtained by a computationally inexpensive model. This model considers the weighted sum of normally distributed *OH adsorption energies with means given by the pure metal constituents of the HEA and with standard deviations set to 0.13 eV as observed previously for *OH adsorption energy distributions on a comparable HEA.[17] The molar fractions of each element in the compositions found with EDX of each HEA constituted the weights of each of the normal distributions in the linear combination given by Equation 2. Here, P is the resulting distribution of *OH adsorption energies, $\Delta G_{^*\mathrm{OH}}$, for a given HEA composition. f_m is the molar fraction of metal m, μ_m is the adsorption energy obtained for the pure metal m, and $\sigma = 0.13$ eV is the

spread in the adsorption energies.

$$
\begin{aligned}
P(\Delta G_{*\mathrm{OH}}) &= \sum_{m}^{metals} f_m \mathcal{N}(\Delta G_{*\mathrm{OH}}; \mu_m, \sigma) \\
&= \sum_{m}^{metals} f_m \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{\Delta G_{*\mathrm{OH}} - \mu_m}{\sigma}\right)^2}
\end{aligned}
\tag{2}
$$

By picking the adsorption energy at the 5 percentile of most strongly adsorbing sites in the modeled *OH adsorption energy distributions for each of the sampled HEA compositions, the trend in Figure 6 is obtained. A somewhat linear relationship can be observed with a capability of predicting onset potentials with an MAE of 11 mV, comparable to the machine learning models. This observed trend hints that the oxidation of water at the catalyst surface and subsequent formation of oxygen adsorbed intermediates could be involved in the potential-limiting step.

Although the *OH adsorption energy seems to explain the overall trend for lowering the onset potential across a large span of *OH adsorption energies, there is not much explained variance for the sampled Ru-rich compositions with strong *OH adsorption at around -0.1 eV. Explaining the subtle differences between the Ru-rich compositions is out of the scope of the current study. However, deeper understanding of the CO oxidation mechanism on randomly populated multi-element surfaces might hold the key for better modeling on HEA catalysts.

Conclusion

In this research, the composition of the PtRuPdRhAu HEA was optimized for the CO oxidation reaction using Bayesian optimization. This provided a guided search among the vast amount of different HEA to find the best performing composition. This optimization required minimal characterization of the nanoparticle catalyst and relied mainly on the synthesis variables and a measure for the catalytic performance. The most active alloy was found within the first optimization cycle, after 32 experiments. Subsequent cycles provided

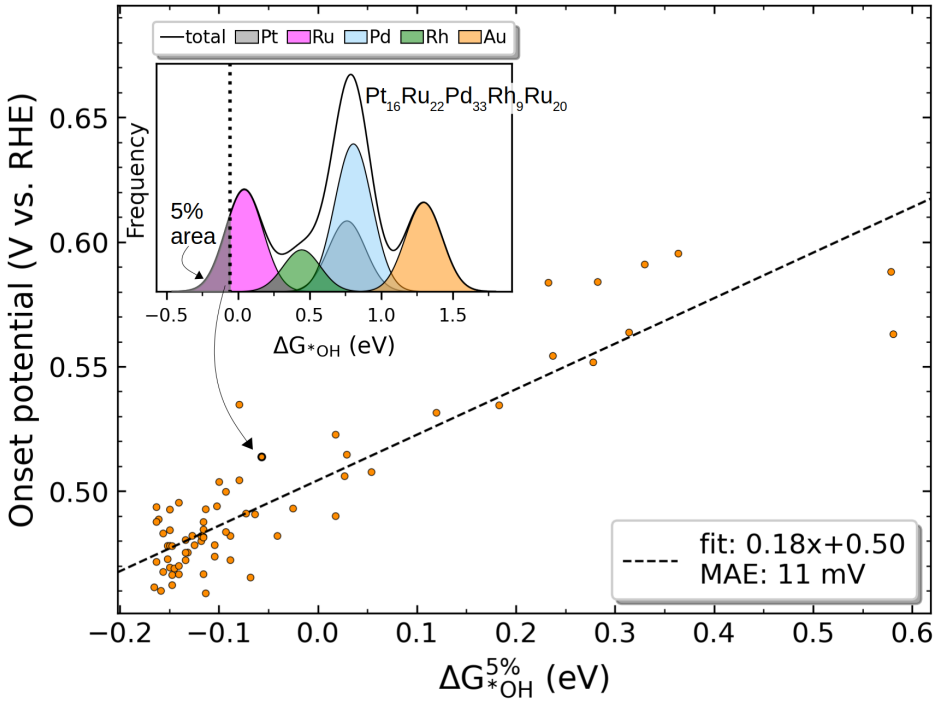

Figure 6: Correlation of experimental onset potentials with the *OH adsorption energy at the 5 percentile of most strongly bound sites of the *OH adsorption energy distribution for the EDX-analyzed compositions. The inset shows an example of the modeled *OH adsorption energy distribution for an HEA sample. The 5 percentile *OH adsorption energy is found as the energy at which 5% of the area of the distribution is to the left.

mostly alloys within the high performance region, with only several exploration points in the low performance region.

The obtained dataset was used to gather knowledge on the CO oxidation reaction and the elemental composition. This was achieved by constructing a random forest regression and a lasso regression. Both models pointed out that Ru is the key element in the CO oxidation reaction. With DFT calculations these relationships were also shown to correlate with the *OH adsorption energy, which is the lowest for Ru. Therefore, within this compositional space, large amounts of Ru are obligatory for a high performing CO oxidation catalyst.

Acknowledgements

This work was supported by the Swiss National Science Foundation (SNSF) via the project No. 200021_184742 and the Danish National Research Foundation Center for High Entropy Alloys Catalysis (CHEAC) DNRF-149. J.P. and K. M. Ø. J. acknowledges support from the Danish Ministry of Higher Education and Science (Structure of Materials in Real Time (SMART) grant). K. M. Ø. J acknowledges support from the Villum Foundation (VKR00015416). S. B. Simonsen and L. Theil Kuhn, Technical University of Denmark are thanked for access to TEM microscopy.

Experimental

Synthesis

The solid precursors, H_2PtCl_6 (Sigma Aldrich, 99.9%) $HAuCl_3$ (Alfa Aesar, 99.99%) $RuCl_3$ (Sigma Aldrich, ReagentPlus), $PdCl_2$ (Sigma Aldrich, ReagentPlus, 99%), and $RhCl_3$ (Sigma Aldrich, 99.98%) were dissolved in methanol (Merck, EMSURE) to produce 20 mM precursor solution. A total of 1 mL of precursor solutions were added in stoichiometric ratios to 5.6 mL 57 mM NaOH MeOH with 50 wt% (metal to carbon) dispersed carbon flakes (Vulkan). The obtained solution was stirred for ten minutes at room temperature and stored overnight in

a centrifuge tube. The following day, the suspensions were centrifuged, decanted and left to dry in air.

Ink were prepared by adding isopropanol and H_2O (3:1, v/v) to the dry catalyst to produce a 0.83 mg metal per mL ink. These inks were drop casted on a glassy carbon RDE tip to produce a metal loading of 35 µg cm^{-2}.

Electrochemistry

All experiments were carried out in a three electrode RDE cell. In between experiments, the cell and all glass components were stored in 1 g mL^{-1} $KMnO_4$ acidified with H_2SO_4. Before the experiments, this solution is was removed and residual MnO_2 was dissolved by adding a dilution of H_2SO_4 and H_2O_2. Following, the cell and components were boiled three times in MilliQ water.

All measurements were carried out with a Nordic Potentiostat EC200. As a reference electrode a trapped RHE electrode was employed. A platinum wire separated by a glass frit was used as a counter electrode. The electrolyte was 0.1 M H_2SO_4 (Merck Suprapur). The working electrode consisted of a Radiometer RDE which has a glassy carbon disk on which the catalyst containing ink was drop casted. This working electrode was put under controlled potential of 0.05 V vs. RHE into the solution. In the first 30 seconds CO gas (Air Liquide, Quality 37) was bubbled through the electrolyte which were followed by 10 cyclic voltammograms (CVs) between 0 and 0.6 V vs. RHE at a speed of 100 mV s^{-1}. After, the electrode was kept for 2 minutes longer at 0.05 V vs. RHE in CO atmosphere. Next, the solution was degassed for 20 minutes with H_2 (ALPHAGAZ, 99.999%) and 2 CVs were recorded between 0 and 1.4 V vs. RHE at a scan speed of 10 mV s^{-1}.

EDX and TEM

The samples were prepared by dropcasting several times 10 µL on copper foil over each other, to yield a think catalyst film. The EDX measurement were performed in a GeminiSEM450

(Zeiss), operated with Smart SEM 6.05, using the EDS Photodetector Ultim max 65 (Oxford instruments), operated with AZTec 4.2. The EDX spectra were measured at four different locations with a size of 800 µm^2 at a working distance of 8 mm and an accelerating voltage of 25 kV.

TEM samples were prepared by dropcasting once the ink on a copper grid. TEM micrographs and their associated EDX spectra were acquired on a JEOL 2100 TEM operated at 200 kV.

Machine learning

Dataset construction

The dataset for the different PtRuRhPdAu alloy compositions was constructed in a Bayesian manner similar to the research of Nugraha et al.[24] The first 25 data points were selected randomly but they included the 5 extreme compositions that span the domain of interest. These extreme HEA compositions consisted of 96% of one element and 1% of each remaining element. The electrochemistry for the CO oxidation was evaluated using the hydrogen oxidation as a probe reaction. Subsequently, only particles that remained active for the HOR after the CO got oxidized were included into machine learning. This made the initial dataset contain a total of 21 points. The input parameter for the machine learning models were the precursor ratios used in the synthesis. The output parameter was the CO oxidation onset potential. This potential was defined at a current increase of 1.5 mA mg^{-1} on top of the capacitive current. For every synthesized alloy, 3 electrochemical experiments were carried out.

This obtained information was used to construct a random forest regression model using the python package scikit learn.[48] The parameters, n_estimatros, bootstrap, max_depth, max_features, min_sample_leaf, min_sample_split, criterion, were tuned using a 500 Bayesian optimization cycles which optimized the k-fold cross validation score. All Bayesian Optimization loops made use of the GpyOpt library in python.[49] The obtained random forest regres-

sion model was further used to find new alloy compositions of interest for investigation. It was optimized using Bayesian optimization, in which the first iteration consisted of the compositions that were used to train the model initially. This way, the Bayesian optimization constructed the prior using the same data that was available in the lab. Following, the Bayesian optimization was continued for an additional 50 iterations, in which new compositions were suggested for evaluation. Among these 50 compositions, a selection of 10 was made based on their distance to other samples. These 10 samples, were synthesized and subsequently added to this optimization cycle. The expansion of the dataset was halted after 5 optimization cycles, as the MAE of the LOOCV showed to reach a constant value around 13 mV. This suggested that an exponential amount of data is becomming necessary to improve the models. In addition, the target of finding the best performing alloy was reached in the first cycle. In the end, this resulted in a dataset of 68 unique catalysts. The scripts used in this paper are available on `https://github.com/vamints/Scripts_BayesOpt_PtRuPdRhAu_paper`.

After the catalysts were analyzed with EDX, new input parameters were obtained, which corresponded to elemental ratios observed by EDX. These allowed to construct a new random forest model, which used the same hyperparameter tuning script as the synthesis model. All random tree models were explained using the SHAP package.[43]

Linear models

Linear models were constructed using the scikit learn package. Features were expanded to include also polynomial terms up to the second degree. The best Lasso Regression model was selected by using a custom loop that varied the lambda penalty until the change in lambda was less than 1%.

The ternary contour plots were created by using the plotly library.[50] The models were fitted using the entire dataset of 68 experimental points. Following these models were used to predict the values of a grid, that spanned the slice of the hyperspace that is shown in the contour plot.

DFT

DFT calculations were done with the GPAW code[51,52] version 19.8.1 and the revised Perdew-Burke-Ernzerhof (RPBE) exchange-correlation functional.[53] Manipulation of atomic structures was performed with the Atomic Simulation Environment (ASE).[54] Four-layered face-centered cubic (fcc) (111) surface slabs measuring 2x2 atoms laterally, and periodically repeated laterally from orthogonal unit cells, were constructed for each of the contituent elements Au, Pd, Pt, Rh, and Ru for *OH adsorption energy calculations. The surface slabs were constructed with fcc lattice constants that were obtained as the minimum energy lattice parameters for the pure fcc bulk constituents (see Table 1).

Table 1: Fcc lattice constants and *OH adsorption free energies of the pure elements used.

Metal	Au	Pd	Pt	Rh	Ru
DFT fcc lattice constant (Å)	4.2149	3.9814	3.9936	3.8648	3.8285
DFT $\Delta G_{*\mathrm{OH}}$ (eV)	1.30	0.80	0.76	0.44	0.04

During structure relaxations all but the two top layers of the slab were fixed, and the *OH adsorbate was put at on-top positions with the oxygen atom contrained to move only perpendicular to the surface. The slabs were constructed with a vacuum of 15 Å above and below the structure. The wave functions were expanded in plane waves with an energy cut-off of 500 eV, and sampling of the Brillouin zone was done on a Monhorst-pack grid of k-points sized 8x8x1. The structures were relaxed so that the maximum force on any atom did not exceed 0.05 eV/Å. For other parameters, the default parameters of GPAW were used.

*OH adsorption free energies were calculated relative to the *OH adsorption energy on Pt(111) which has been shown to adsorb *OH about 0.1 eV stronger[55] than the maximum of the oxygen reduction reaction (ORR) activity volcano at 0.86 eV relative to H_2O(l) and H_2(g),[56] i.e. setting the free energy of *OH adsorption on Pt(111) to about 0.76 eV.

$$\Delta \mathrm{G}_{*\mathrm{OH}} = (\mathrm{E}_{*\mathrm{OH}} - \mathrm{E}_{*}) - \left(\mathrm{E}_{*\mathrm{OH}}^{\mathrm{Pt(111)}} - \mathrm{E}_{*}^{\mathrm{Pt(111)}}\right) + 0.76\,\mathrm{eV} \tag{3}$$

Here, $\Delta \mathrm{G}_{*\mathrm{OH}}$ is the free energy of *OH adsorption, $\mathrm{E}_{*\mathrm{OH}}$ and E_{*} are the DFT calculated

energies of the surface slab with and without *OH adsorbed, respectively. The adsorption energies used in this work are given in Table 1.

References

(1) Andújar, J. M.; Segura, F. Fuel cells: History and updating. A walk along two centuries. *Renewable and Sustainable Energy Reviews* **2009**, *13*, 2309–2322.

(2) Rath, R.; Kumar, P.; Mohanty, S.; Nayak, S. K. Recent advances, unsolved deficiencies, and future perspectives of hydrogen fuel cells in transportation and portable sectors. *International Journal of Energy Research* **2019**, *43*, 8931–8955.

(3) Alaswad, A.; Omran, A.; Sodre, J. R.; Wilberforce, T.; Pignatelli, G.; Dassisti, M.; Baroutaji, A.; Olabi, A. G. Technical and Commercial Challenges of Proton-Exchange Membrane (PEM) Fuel Cells. *Energies* **2020**, *14*, 144.

(4) Gasteiger, H. A.; Kocha, S. S.; Sompalli, B.; Wagner, F. T. Activity benchmarks and requirements for Pt, Pt-alloy, and non-Pt oxygen reduction catalysts for PEMFCs. *Applied Catalysis B: Environmental* **2005**, *56*, 9–35.

(5) Gasteiger, H. A.; Panels, J. E.; Yan, S. G. Dependence of PEM fuel cell performance on catalyst loading. *Journal of Power Sources* **2004**, *127*, 162–171.

(6) Jackson, C.; Raymakers, L. F.; Mulder, M. J.; Kucernak, A. R. Poison mitigation strategies for the use of impure hydrogen in electrochemical hydrogen pumps and fuel cells. *Journal of Power Sources* **2020**, *472*, 228476.

(7) Zheng, T.; Zhou, W.; Li, X.; You, H.; Yang, Y.; Yu, W.; Zhang, C.; Chu, X.; San Hui, K.; Ding, W. Structural design of self-thermal methanol steam reforming microreactor with porous combustion reaction support for hydrogen production. *International Journal of Hydrogen Energy* **2020**, *45*, 22437–22447.

(8) Gurau, V.; Ogunleke, A.; Strickland, F. Design of a methanol reformer for on-board production of hydrogen as fuel for a 3 kW High-Temperature Proton Exchange Membrane Fuel Cell power system. *International Journal of Hydrogen Energy* **2020**, *45*, 31745–31759.

(9) Uchida, H.; Izumi, K.; Watanabe, M. Temperature dependence of CO-tolerant hydrogen oxidation reaction activity at Pt, Pt-Co, and Pt-Ru electrodes. *Journal of Physical Chemistry B* **2006**, *110*, 21924–21930.

(10) Koper, M. T. Electrocatalysis on bimetallic and alloy surfaces. *Surface Science* **2004**, *548*, 1–3.

(11) Freitas, K. S.; Lopes, P. P.; Ticianelli, E. A. Electrocatalysis of the hydrogen oxidation in the presence of CO on RhO2/C-supported Pt nanoparticles. *Electrochimica Acta* **2010**, *56*, 418–426.

(12) Watanabe, M.; Furuuchi, Y.; Motoo, S. Electrocatalysis by AD-atoms. *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry* **1985**, *191*, 367–375.

(13) Buatier de Mongeot, F.; Scherer, M.; Gleich, B.; Kopatzki, E.; Behm, R. CO adsorption and oxidation on bimetallic Pt/Ru(0001) surfaces – a combined STM and TPD/TPR study. *Surface Science* **1998**, *411*, 249–262.

(14) Yeh, J.-W.; Chen, S.-K.; Lin, S.-J.; Gan, J.-Y.; Chin, T.-S.; Shun, T.-T.; Tsau, C.-H.; Chang, S.-Y. Nanostructured High-Entropy Alloys with Multiple Principal Elements: Novel Alloy Design Concepts and Outcomes. *Advanced Engineering Materials* **2004**, *6*, 299–303.

(15) Zou, Y.; Ma, H.; Spolenak, R. Ultrastrong ductile and stable high-entropy alloys at small scales. *Nature Communications* **2015**, *6*, 7748.

(16) Ye, Y.; Wang, Q.; Lu, J.; Liu, C.; Yang, Y. High-entropy alloy: challenges and prospects. *Materials Today* **2016**, *19*, 349–362.

(17) Batchelor, T. A.; Pedersen, J. K.; Winther, S. H.; Castelli, I. E.; Jacobsen, K. W.; Rossmeisl, J. High-Entropy Alloys as a Discovery Platform for Electrocatalysis. *Joule* **2019**, *3*, 834–845.

(18) Xin, Y.; Li, S.; Qian, Y.; Zhu, W.; Yuan, H.; Jiang, P.; Guo, R.; Wang, L. High-Entropy Alloys as a Platform for Catalysis: Progress, Challenges, and Opportunities. *ACS Catalysis* **2020**, *10*, 11280–11306.

(19) Löffler, T.; Savan, A.; Garzón-Manjón, A.; Meischein, M.; Scheu, C.; Ludwig, A.; Schuhmann, W. Toward a Paradigm Shift in Electrocatalysis Using Complex Solid Solution Nanoparticles. *ACS Energy Letters* **2019**, *4*, 1206–1214.

(20) Yao, Y. et al. Computationally aided, entropy-driven synthesis of highly efficient and durable multi-elemental alloy catalysts. *Science Advances* **2020**, *6*, eaaz0510.

(21) Ludwig, A. Discovery of new materials using combinatorial synthesis and high-throughput characterization of thin-film materials libraries combined with computational methods. *npj Computational Materials* **2019**, *5*, 70.

(22) Batchelor, T. A. A.; Löffler, T.; Xiao, B.; Krysiak, O. A.; Strotkötter, V.; Pedersen, J. K.; Clausen, C. M.; Savan, A.; Li, Y.; Schuhmann, W.; Rossmeisl, J.; Ludwig, A. Complex-Solid-Solution Electrocatalyst Discovery by Computational Prediction and High-Throughput Experimentation**. *Angewandte Chemie International Edition* **2021**, *60*, 6932–6937.

(23) Pedersen, J. K.; Batchelor, T. A.; Bagger, A.; Rossmeisl, J. High-entropy alloys as catalysts for the CO_2 and CO reduction reactions. *ACS Catalysis* **2020**, *10*, 2169–2176.

(24) Nugraha, A. S.; Lambard, G.; Na, J.; Hossain, M. S. A.; Asahi, T.; Chaikittisilp, W.; Yamauchi, Y. Mesoporous trimetallic PtPdAu alloy films toward enhanced electrocatalytic activity in methanol oxidation: Unexpected chemical compositions discovered by Bayesian optimization. *Journal of Materials Chemistry A* **2020**, *8*, 13532–13540.

(25) Pedersen, J. K.; Clausen, C. M.; Krysiak, O. A.; Xiao, B.; Batchelor, T. A. A.; Löffler, T.; Mints, V. A.; Banko, L.; Arenz, M.; Savan, A.; Schuhmann, W.; Ludwig, A.; Rossmeisl, J. Bayesian Optimization of High-Entropy Alloy Compositions for Electrocatalytic Oxygen Reduction. `https://arxiv.org/abs/2106.08212`, 2021.

(26) Mistry, A.; Franco, A. A.; Cooper, S. J.; Roberts, S. A.; Viswanathan, V. How Machine Learning Will Revolutionize Electrochemical Sciences. *ACS Energy Letters* **2021**, 1422–1431.

(27) Xie, Y.; Zou, J.; Li, Z.; Gao, F.; Peng, C. A novel deep belief network and extreme learning machine based performance degradation prediction method for proton exchange membrane fuel cell. *IEEE Access* **2020**, *8*, 176661–176675.

(28) Wang, F. K.; Huang, C. Y.; Mamo, T.; Cheng, X. B. Ensemble model for the degradation prediction of proton exchange membrane fuel cell stacks. *Quality and Reliability Engineering International* **2021**, *37*, 34–46.

(29) Min, K.; Choi, B.; Park, K.; Cho, E. Machine learning assisted optimization of electrochemical properties for Ni-rich cathode materials. *Scientific Reports* **2018**, *8*, 1–7.

(30) Wang, M.; Wang, T.; Cai, P.; Chen, X. Nanomaterials Discovery and Design through Machine Learning. *Small Methods* **2019**, *3*, 1–7.

(31) Gasteiger, H. A.; Markovic, N.; Ross, P. N.; Cairns, E. J. CO electrooxidation on well-characterized Pt-Ru alloys. *Journal of physical chemistry* **1994**, *98*, 617–625.

(32) Rodriguez, P.; Koper, M. T. Electrocatalysis on gold. *Physical Chemistry Chemical Physics* **2014**, *16*, 13583–13594.

(33) Tian, M.; Conway, B. E. Phenomenology of oscillatory electro-oxidation of formic acid at Pd: Role of surface oxide films studied by voltammetry, impedance spectroscopy and nanogravimetry. *Journal of Electroanalytical Chemistry* **2005**, *581*, 176–189.

(34) Lai, S. C.; Lebedeva, N. P.; Housmans, T. H.; Koper, M. T. Mechanisms of carbon monoxide and methanol oxidation at single-crystal electrodes. *Topics in Catalysis* **2007**, *46*, 320–333.

(35) Nibbelke, R. H.; Nievergeld, A. J.; Hoebink, J. H.; Marin, G. B. Development of a transient kinetic model for the CO oxidation by O_2 over a $Pt/Rh/CeO_2/\gamma$-Al_2O_3 three-way catalyst. *Applied Catalysis B: Environmental* **1998**, *19*, 245–259.

(36) González-Velasco, J. R.; Botas, J. A.; Ferret, R.; Pilar González-Marcos, M.; Marc, J. L.; Gutiérrez-Ortiz, M. A. Thermal aging of Pd/Pt/Rh automotive catalysts under a cycled oxidizing-reducing environment. *Catalysis Today* **2000**, *59*, 395–402.

(37) Quinson, J.; Neumann, S.; Wannmacher, T.; Kacenauskaite, L.; Inaba, M.; Bucher, J.; Bizzotto, F.; Simonsen, S. B.; Theil Kuhn, L.; Bujak, D.; Zana, A.; Arenz, M.; Kunz, S. Colloids for Catalysts: A Concept for the Preparation of Superior Catalysts of Industrial Relevance. *Angewandte Chemie - International Edition* **2018**, *57*, 12338–12341.

(38) Quinson, J.; Kacenauskaite, L.; Bucher, J.; Simonsen, S. B.; Theil Kuhn, L.; Oezaslan, M.; Kunz, S.; Arenz, M. Controlled Synthesis of Surfactant-Free Water-Dispersible Colloidal Platinum Nanoparticles by the Co4Cat Process. *ChemSusChem* **2019**, *12*, 1229–1239.

(39) Quinson, J.; Neumann, S.; Kacenauskaite, L.; Bucher, J.; Kirkensgaard, J. J.; Simonsen, S. B.; Theil Kuhn, L.; Zana, A.; Vosch, T.; Oezaslan, M.; Kunz, S.; Arenz, M.

Solvent-Dependent Growth and Stabilization Mechanisms of Surfactant-Free Colloidal Pt Nanoparticles. *Chemistry - A European Journal* **2020**, *26*, 9012–9023.

(40) Quinn, M.; Mills, G. Surface-mediated formation of gold particles in basic methanol. *Journal of Physical Chemistry* **1994**, *98*, 9840–9844.

(41) Quinson, J.; Simonsen, S. B.; Kuhn, L. T.; Kunz, S.; Arenz, M. Size effect studies in catalysis: a simple surfactant-free synthesis of sub 3 nm Pd nanocatalysts supported on carbon. *RSC Advances* **2018**, *8*, 33794–33797.

(42) James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *Linear Model Selection and Regularization*; Springer Texts in Statistics; Springer New York: New York, NY, pp 203–264.

(43) Lundberg, S. M.; Lee, S.-I. In *Advances in Neural Information Processing Systems 30*; Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc., 2017; pp 4765–4774.

(44) Lundberg, S. M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J. M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* **2020**, *2*, 2522–5839.

(45) Papandrew, A. B.; Atkinson, R. W.; Unocic, R. R.; Zawodzinski, T. A. Ruthenium as a CO-tolerant hydrogen oxidation catalyst for solid acid fuel cells. *Journal of Materials Chemistry A* **2015**, *3*, 3984–3987.

(46) Goyal, A.; Marcandalli, G.; Mints, V. A.; Koper, M. T. M. Competition between CO_2 Reduction and Hydrogen Evolution on a Gold Electrode under Well-Defined Mass Transport Conditions. *Journal of the American Chemical Society* **2020**, *142*, 4154–4161.

(47) Lin, W. F.; Zei, M. S.; Eiswirth, M.; Ertl, G.; Iwasita, T.; Vielstich, W. Electrocatalytic

activity of Ru-modified Pt(111) electrodes toward CO oxidation. *Journal of Physical Chemistry B* **1999**, *103*, 6968–6977.

(48) Pedregosa, F. et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.

(49) The GPyOpt authors, GPyOpt: A Bayesian Optimization framework in python. `http://github.com/SheffieldML/GPyOpt`, 2016.

(50) Plotly Technologies Inc., Collaborative data science. 2015; `https://plot.ly`.

(51) Mortensen, J. J.; Hansen, L. B.; Jacobsen, K. W. Real-space grid implementation of the projector augmented wave method. *Phys. Rev. B* **2005**, *71*, 035109.

(52) Enkovaara, J. et al. Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *Journal of Physics: Condensed Matter* **2010**, *22*, 253202.

(53) Hammer, B.; Hansen, L. B.; Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **1999**, *59*, 7413–7421.

(54) Larsen, A. H. et al. The atomic simulation environment—a Python library for working with atoms. *Journal of Physics: Condensed Matter* **2017**, *29*, 273002.

(55) Stephens, I. E. L.; Bondarenko, A. S.; Grønbjerg, U.; Rossmeisl, J.; Chorkendorff, I. Understanding the electrocatalysis of oxygen reduction on platinum and its alloys. *Energy Environ. Sci.* **2012**, *5*, 6744–6762.

(56) Pedersen, J. K.; Batchelor, T. A.; Yan, D.; Skjegstad, L. E. J.; Rossmeisl, J. Surface electrocatalysis on high-entropy alloys. *Current Opinion in Electrochemistry* **2021**, *26*, 100651.

Supporting Information

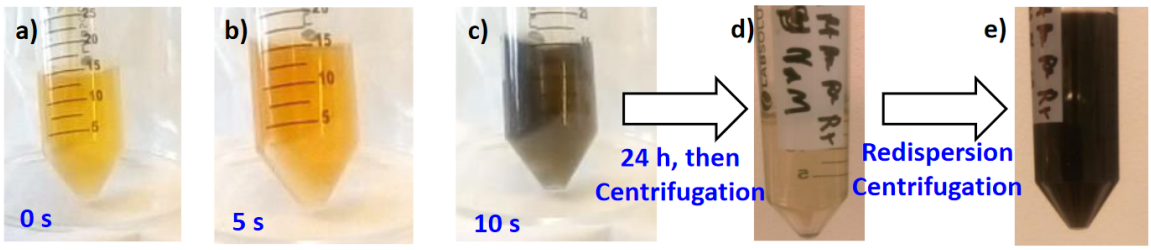

Figure S1: Time series photographs of H_2PtCl_6, $IrCl_3$, $RuCl_3$ and $PdCl_2$ in alkaline NaOH methanol under stirring at room temperature **a)** before, **b)** 5 s and **c)** 10 s after addition of $HAuCl_4$. **d)** is the same solution after 24 hours of synthesis and after centrifugation whereas **e)** is the same solution as **d)** after removing the supernatant, redispersion in methanol and a new centrifugation cycle. This reaction is for illustration purposes of the simplicity of the general synthesis approach and no carbon support was used to more clearly obverse the speed of the reaction (in presence of support the reaction mixture is completely black). The final metal concentration for each element is here 0.5 mM and the final NaOH concentration is here 48 mM.

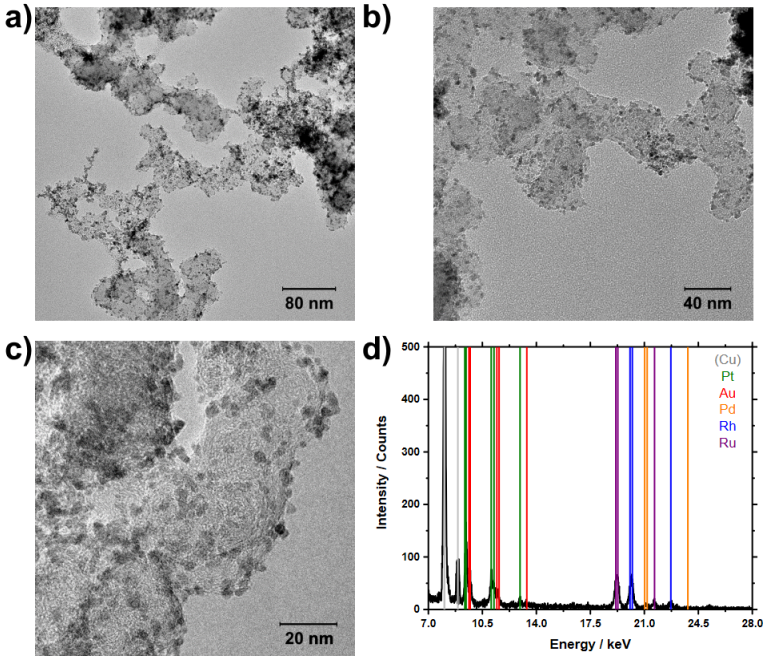

Figure S2: **a-c)** TEM and **d)** EDX characterization of the sample $Pt_{12}Ru_{38}Pd_8Rh_{41}Au_1$ with the expected intensity peaks positions for different elements as indicated in colour in **d)** where the spectrum is in dark. The copper signal (Cu) comes from the TEM grid.

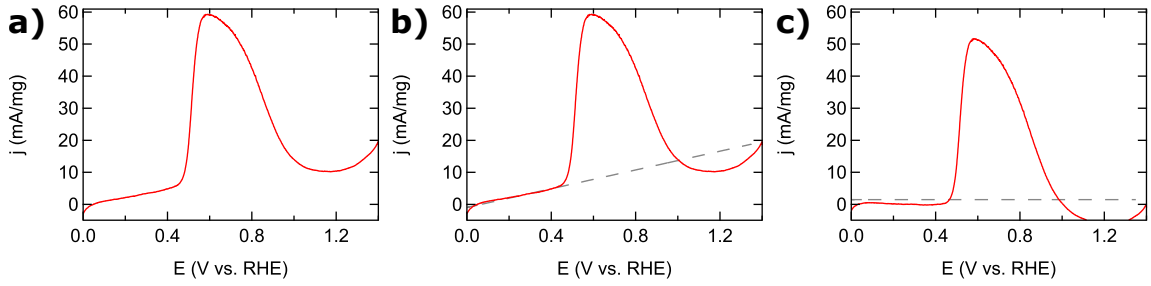

Figure S3: Example of how the data was obtained. **a)** Raw data as measured experimentally. **b)** Dashed line shows the linear fit through the capacitance. **c)** Dashed line indicates the current density at which the CO oxidation onset potential is defined.

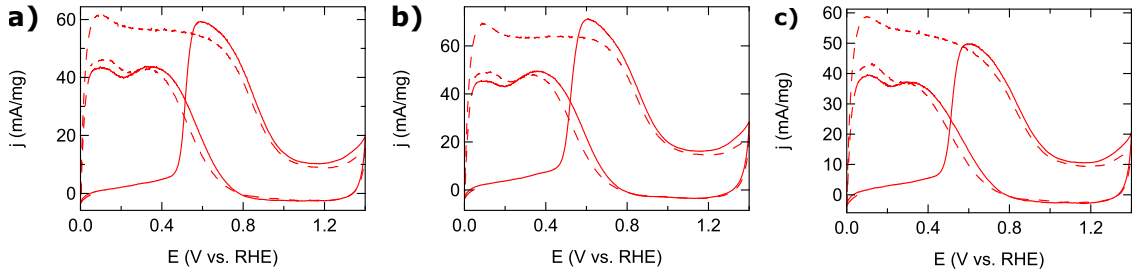

Figure S4: The CO monolayer oxidation on a 35 µg cm^{-2} $Pt_{12}Ru_{38}Pd_8Rh_{41}Au_1$ film. **a-c** CVs were recorded in H_2 atm after a CO monolayer was adsorbed on the surfaces at 10 mV s^{-1} from 0 to 1.4 V vs. RHE. The solid line represents the first cycle, the dashed line represents the second cycle.

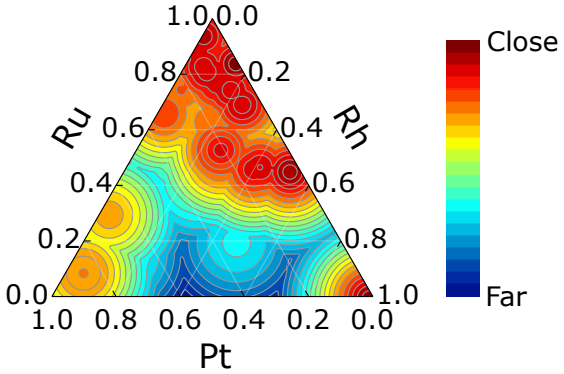

Figure S5: The distance of the grids points for the EDX models in the ternary plots (Figure 5a-b) of Ru, Pt, Rh to the closest measured experimental data point in the five dimensional space of RuPtRhPdAu.