# Beyond Scaling Relations for the Description of Catalytic Materials

Mie Andersen,*[,†] Sergey V. Levchenko,[‡,⊥] Matthias Scheffler,[‡] and Karsten Reuter[†]

[†]Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, Lichtenbergstrasse 4, 85747 Garching, Germany

[‡]Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

Ⓢ *Supporting Information*

ABSTRACT: Computational screening for new and improved catalyst materials relies on accurate and low-cost predictions of key parameters such as adsorption energies. Here, we use recently developed compressed sensing methods to identify descriptors whose predictive power extends over a wide range of adsorbates, multimetallic transition metal surfaces, and facets. The descriptors are expressed as nonlinear functions of intrinsic properties of the clean catalyst surface, e.g. coordination numbers, d-band moments, and density of states at the Fermi level. From a single density functional theory calculation of these properties, we predict adsorption energies at all potential surface sites, and thereby also the most stable geometry. Compared to previous approaches such as scaling relations, we find our approach to be both more general and more accurate for the prediction of adsorption energies on alloys with mixed-metal surfaces, already when based on training data including only pure metals. This accuracy can be systematically improved by also adding alloy adsorption energies to the training data.

physically motivated descriptors   data-driven descriptors
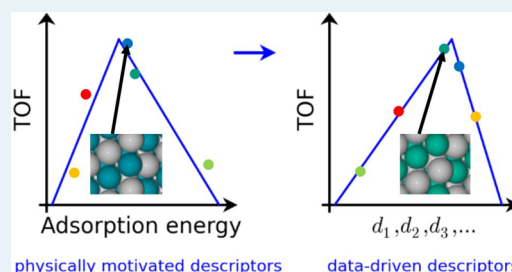
KEYWORDS: computational screening, catalyst, adsorption energies, density functional theory, descriptor

INTRODUCTION

Surface catalysis has an enormous impact on the environment and our society's health and prosperity.[1] However, the reliable description of catalytic properties and the prediction of what materials may be even better catalysts than what we know today is still weak. This is due to the highly nonlinear and intricate relationship between the "catalyst material" (the static material that is introduced before the catalytic process is running) and the strongly kinetically controlled surface reactions at realistic conditions.[2,3] Simply speaking, the basic understanding of heterogeneous catalysis is given by the Sabatier principle and the Brønsted-Evans−Polanyi (BEP) relation.[4] The first states that there is an optimum adsorption strength for which the reactants bind strong enough to allow for adsorption and dissociation into reaction intermediates, but weak enough to allow for consecutive desorption of products. In turn, the BEP concept tells that the energy barriers of the chemical reactions scale approximately linearly with the adsorption energies of the molecules.

In consequence, the reliable prediction of adsorption energies is a key element of any theoretical description and search for new catalyst materials. For this, we here present a data-driven approach that does not start from a specific physical model, e.g. the tight-binding description of chemical bonding, but accepts that the intricacy of processes that cooperate or compete in materials properties may not necessarily be describable by a closed physical equation. This has been described as the fourth paradigm of materials science.[5] Previous data-driven approaches to the prediction of

adsorption energies[6−12] have exploited an approximately linear correlation between the adsorption energies of certain adsorbates on pure transition metal (TM) surfaces (scaling relations[13]) to extend the data-driven predictions of one or two species to other adsorbates involved in the reaction. We instead directly learn the adsorption energies of a whole range of atoms and molecules at all potential adsorption sites (thereby also the most stable site) only from properties of the clean surface. That is, our approach allows us to learn the adsorption energies of all species involved in a given reaction network with no assumptions of linearity. We thereby go beyond scaling relations and the often unfulfilled assumptions tied to this particular physical model. This furthermore opens the perspective of directly searching for outliers to scaling relations, which can be highly interesting catalyst materials missed by the standard approach.[14−16]

The method for identifying the key descriptive parameters is the recently developed compressed sensing method SISSO[17] (sure independence screening and sparsifying operator), which enables us to identify the best multidimensional descriptor out of an immensity of candidates (billions). The simultaneous learning of the adsorption energies of a whole range of atoms and molecules is made possible by a multitask learning approach,[18] wherein one single descriptor that works for all considered atoms and molecules is identified. Our descriptors

are more general and less costly to use than previous approaches and allow for making predictions for a huge number of surfaces, including both multimetallics and various facets. Through BEP relations, our approach can also describe chemical reactions as well as the diffusion of atoms and molecules at the surface visiting metastable adsorption sites.

■ COMPUTATIONAL DETAILS

Density Functional Theory. The data sets employed in the present work were obtained from plane-wave density functional theory (DFT) calculations (Quantum ESPRESSO code[19]) using the van der Waals-corrected BEEF-vdW exchange-correlation functional.[20] The larger data set consists of adsorption energies of atomic and molecular adsorbates (C, CH, CO, H, O, OH) on the stepped fcc(211) facets of nine TMs (Ni, Cu, Ru, Rh, Pd, Ag, Ir, Pt, and Au) and selected single-atom (SA) and AB bimetallic alloys. The metals were modeled in fcc stacking using a (1×3) $((1 \times 2)$ for AB alloys) supercell with 15 metal layers, which corresponds to five layers along the (111) direction. The considered adsorption sites are illustrated in Figure 1(a) and cover both high-symmetry
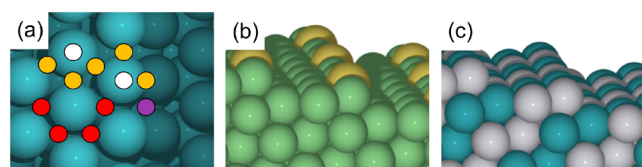
Figure 1. Top view of the structure of (a) the fcc(211) facet along with the considered terrace and step adsorption sites, which cover one fourfold-coordinated site (purple dot), four threefold-coordinated (fcc and hcp) sites (red dots), five bridge sites (yellow dots), and two top sites (white dots). Perspective views of the structures of (b) a single-atom (SA) alloy and (c) an AB alloy.

terrace and step sites. For each adsorbate, all adsorption sites that correspond to local minima on the potential energy surface were included. The SA alloys[15,16,21] were constructed by replacing one metal atom at the step with a different metal (see Figure 1(b)). Specifically, we considered Ag@Cu (Ag atom in Cu surface), Pt@Rh, Pd@Ir, and Au@Ni. For the AB L1$_0$ alloys[22] (AgPd, IrRu, PtRh, and AgAu), the considered surface termination is depicted in Figure 1(c). The total numbers of adsorption energies in the data set are 344 (metals), 281 (SA alloys), and 259 (AB alloys). The smaller facets data set consists of one adsorption site for each adsorbate on the fcc(111), (110), and (100) facets of the 9 TMs, leading to a total of 54 adsorption energies on each facet. All data sets are compiled in Supplementary Section S1 together with further computational details.[23]

Compressed Sensing. The SISSO method[17] employed for descriptor identification makes the ansatz that the properties of interest $P_1^j, ..., P_N^j \in \mathbb{R}$, (in this case a vector of N adsorption energies of adsorbate j) can be expressed as linear functions of candidate features $\mathbf{d}_1, ..., \mathbf{d}_M \in \mathbb{R}^N$, where the features are constructed as nonlinear functions of user-defined primary features (see below). SISSO identifies the few best features (the number of which corresponds to the dimensionality of the descriptor) out of immense feature spaces by use of the sparsifying l_0 constraint. This is carried out in a smaller feature subspace selected by a screening procedure (sure independence screening (SIS)). The size of the subspace

is equal to a user-defined SIS value times the dimension of the descriptor.

In this work, we make use of multitask learning[18] to identify common descriptors for the adsorption energies of several different adsorbates simultaneously. That is, the identified features are constrained to be identical for every adsorbate, while the fitting coefficients are allowed to vary between the adsorbates. We find that multitask learning gives a better predictive performance compared to the identification of separate descriptors for each adsorbate (see Supplementary Figure S2). We consider two hyperparameters in the SISSO method: the dimension of the descriptor and the feature space rung (see below) as well as the SIS value (see Supplementary Section S2), that we fix for the current application through a validation data set (see below).

Primary Features. The decision of which primary features to use as input for the feature construction is crucial for the predictive performance of the resulting descriptors. Inspired by previous studies,[6,7,24−31] we consider four classes of primary features (see Table 1) related to the metal atom, metal bulk,

Table 1. Primary Features Used for the Feature Construction

class	name	abbreviation
atomic	Pauling electronegativity	PE
	ionization potential	IP
	electron affinity	EA
bulk	fcc nearest neighbor distance	bulk$_{nnd}$
	radius of d-orbitals	r_d
	coupling matrix element squared	V_{ad}^2
surface	work function	W
site	number of atoms in ensemble	site$_{no}$
	coordination number	CN
	nearest neighbor distance	site$_{nnd}$
	d-band center	ε_d
	d-bandwidth	W_d
	d-band skewness	S_d
	d-band kurtosis	K_d
	d-band filling	f_d
	sp-band filling	f_{sp}
	density of d-states at Fermi level	DOS$_d$
	density of sp-states at Fermi level	DOS$_{sp}$

metal surface, and metal adsorption site. For pure metals, the primary features of the site class were calculated as averages over the metal atoms making up the site ensemble, while for alloys, this was the case for the primary features of all classes. We note that the consideration of fixed adsorption sites as well as the averaging over the site ensemble is an approximation. It may break down in case of surface reconstruction or any other appearance of new adsorption motifs that were not accounted for in the calculation of the primary features. Further details regarding the primary features and all data are given in Supplementary Section S1.

Feature Construction. As discussed above, candidate features are constructed as nonlinear functions of the primary features. In the SISSO method, this is achieved in practice by applying algebraic/functional operators such as addition, multiplication, exponentials, powers, roots, etc. to the features.[17] A full list of the used operators can be found in Supplementary Section S2. Arbitrarily large feature spaces can be constructed by iteratively applying these operators to the
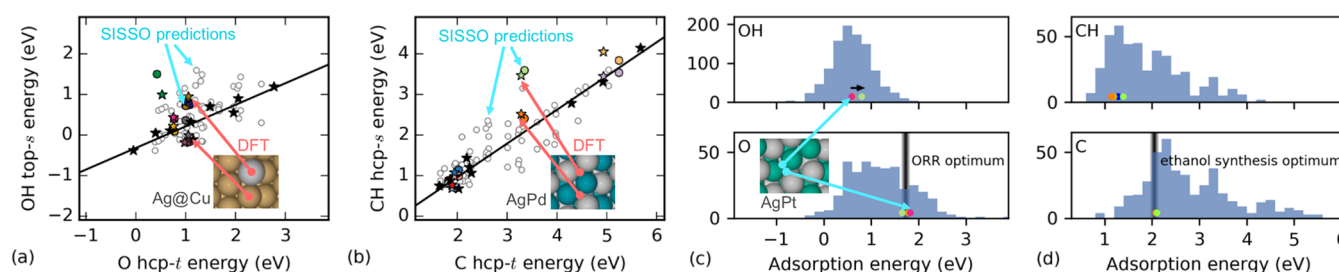
Figure 2. Scaling relations for adsorption energies of (a) OH at the top-s site and (b) CH at the hcp-s site. The black lines show the linear scaling relations fitted to the DFT data on the pure metals (shown as colored stars). Explicitly calculated DFT adsorption energies for (a) SA alloys and (b) AB alloys are shown with colored stars. Some particularly large deviations between predictions from scaling relations and actual adsorption energies for the alloys are highlighted. SISSO predictions (8D, Φ_3 descriptor trained on the pooled metals and alloys data set, see text) for the calculated alloys (32 additional AB alloys) are shown with colored (gray) circles. For the prediction of the DFT-calculated alloys, these predicted data points were excluded from the training set. All scaling relations can be found in Supplementary Figure S3. Histograms of SISSO-predicted adsorption energies on all potential adsorption sites of all 36 AB alloys for (c) O and OH and (d) C and CH. The black shaded regions highlight literature "volcano optimal" adsorption energies for the oxygen reduction reaction (ORR) on (111) facets[36] and for selective ethanol synthesis on (211) facets.[37] The colored circles mark those materials for which the predicted most stable (c) O adsorption energy among the (111)-like (terrace) sites and (d) C adsorption energy among all (211) sites falls within the desired range. The corresponding most stable OH and CH adsorption energies for these materials are also marked in the histograms above. A predicted near-optimal ORR material (AgPt) that breaks the O–OH scaling relation due to different metal compositions of the preferred O and OH adsorption sites is highlighted. The black arrow points from the SISSO-predicted to the scaling-relation-predicted OH adsorption energy.

already generated features. The starting point Φ_0 corresponds to the 18 primary features listed in Table 1. We consider up to three iterations, generating thereby the feature spaces Φ_1, Φ_2, and Φ_3. Note that a given feature space Φ_n also contains all of the lower rung feature spaces. The Φ_1 and Φ_2 feature spaces are still comparatively small. They consist of 783 and about 10^6 features, respectively. For the third iteration generating Φ_3, the approach we chose consisted of carrying out two rounds of feature construction and descriptor identification, each for a subset of only 16 out of the 18 primary features, to limit the Φ_3 feature space of each round to a tractable value of about 10^{11}. In the first round, the skewness and kurtosis of the d-band were excluded because the higher order d-band moments are expected to be less important than the lower order moments. Among the identified best descriptors (i.e., with lowest validation errors, see below) of the first round, two primary features of the site class never appeared, namely the nearest neighbor distance and the density of d-states at the Fermi level. In the second round, these two primary features were then excluded, while the skewness and kurtosis of the d-band were reincluded. At every dimension, the best performing Φ_3 descriptor originated from the first round, and therefore, only the results of the first round are presented below.

RESULTS AND DISCUSSION

Scaling Relations. We begin by evaluating the performance of prevalent scaling relations for predicting adsorption energies on SA and AB alloys. In Figures 2(a) and (b) we show two examples of scaling relations constructed by linear fits to the DFT-calculated adsorption energies on the pure TMs (black stars). Corresponding explicitly calculated adsorption energies on SA and AB alloys are also indicated by colored stars. While many bimetallics are well-described by the linear scaling relations, there are also a number of serious outliers. Some systems with particularly large prediction errors of the order of 1 eV are highlighted. They typically contain mixed-metal sites made up of metals with very different reactivity toward O (e.g., Cu and Ag) or C (e.g., Ag and Pd). This poor performance of scaling relations derives from their calculation of the descriptors at one specific site on the alloy surface.

While scaling relations derived from the pure metals generally do take into account the effect of the different geometries of the various sites on the fcc(211) facet (see Supplementary Figure S3), they fail to account for the variation in metal composition of these sites on alloy surfaces. This issue likely occurs most severely for the considered thermochemical scaling relations. BEP relations for activation energies, in contrast, are more local in the sense that often both the transition state and the initial and final reaction intermediates coordinate (or can be chosen to coordinate) to the same metal atoms at the considered site ensemble. Correspondingly, BEP relations are typically found to exhibit errors significantly lower than those for thermochemical scaling relations even for the pure metals.[32]

We note that an alternative scaling-relation-based approach for alloy screening is to consider only the most stable adsorption sites.[33] Because in general the most stable adsorption sites of the descriptor atoms (e.g., O and C) are not known in advance for an unknown alloy, this entails calculating all potential adsorption sites on the alloy surface. However, at concomitantly increased screening costs, this still does not alleviate the problem because different adsorbates (e.g., O and OH) generally adsorb to different site types (e.g., O typically prefers higher coordinated sites than OH). As a consequence, the metal composition of the preferred sites could be different. We will come back to this point in the discussion of the compressed sensing results. In addition, not only the most stable sites but also metastable sites missed by this approach can get populated at higher coverages and then play an important role in the catalytic pathway.[32,34,35]

Descriptor Identification. The demonstrated failure of scaling relations to predict accurate adsorption energies on alloys with mixed-metal surfaces, as well as the high cost associated with the calculation of two or more adsorption energies on each alloy to be screened, emphasizes the need for new, accurate, and low-cost descriptors for computational screening. In Figure 3(a), we compare the performance of scaling relations to new descriptors identified by SISSO in terms of the root-mean-square error (RMSE) on training and validation data sets. We define the best descriptor as the
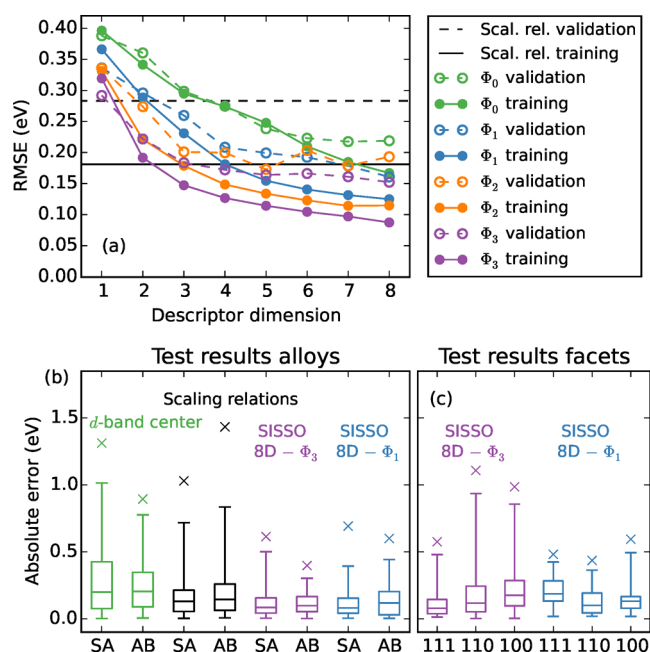
Figure 3. (a) RMSE for the descriptors identified using exclusively the pure metals data set for training and 50% of the alloys data for validation as well as corresponding results for scaling relations. (b) Box plots of the absolute errors on the test set consisting of the remaining 50% of the SA and AB alloys data for the d-band center, for scaling relations, for the best SISSO descriptor identified using the alloys validation data set $(8D, \Phi_3)$, and for the best SISSO descriptor identified when including 50% of the (111), (110), and (100) facets data set in the validation data set $(8D, \Phi_1$, see Supplementary Figure S6). The upper and lower limits of the rectangles mark the 75 and 25% percentiles; the internal horizontal line marks the median, and the "error bars" mark the 99th and 1st percentiles. The crosses mark the maximum absolute errors. (c) Corresponding box plots for the two SISSO descriptors on the facets data sets, where for the $(8D, \Phi_1)$ descriptor, only the remaining 50% of the facets data set not used for validation are included.

descriptor that achieves the lowest RMSE on the validation data set. In the calculation of the RMSE, the same weight is given to every adsorbate considered in the multitask learning irrespective of how many data points exist for the adsorbate (see Supplementary Table S1). The training data consist exclusively of the adsorption energies on the *pure* metals, and the validation data consist of 50% of each of the SA and AB alloy data. SISSO data are shown for 1D−8D descriptors identified from each considered feature space Φ_n. For each case, a number of SIS values have also been tested (see Supplementary Figure S4) and the best descriptor (i.e., the descriptor with the lowest RMSE on the validation data set) is shown. As expected, the SISSO training errors systematically decrease when increasing either the complexity and size of the feature space (larger n) or the dimensionality of the descriptor. The validation errors show the same trend, but the errors level out around the 5D to 8D descriptor depending on the rung of the used feature space. We would expect the validation errors to increase again at even higher dimensions due to overfitting. However, such higher dimensions are outside of the scope of the present study because already the leveling out of the validation errors suggests that going beyond 8D is unlikely to result in descriptors with lower validation errors.

The 5D to 8D descriptors of Φ_3 all have very low validation errors, differing from each other by only about 10 meV. Likely,

there is no statistically significant difference in their performance. A detailed statistical analysis to derive error bars is outside of the scope of this work, which aims at a first comparison to scaling relations. The latter are usually derived based on only one fixed training data set considering only the pure metals, and we therefore follow the same approach here. In the absence of error bars, we choose the descriptor with the lowest observed validation RMSE (of 0.15 eV) as our best descriptor, i.e. the 8D, Φ_3 descriptor. This descriptor (and any other optimized 5D to 8D descriptor) is significantly better than scaling relations, for which the validation RMSE is 0.28 eV (horizontal dashed line). In fact, already the 2D descriptor of Φ_3 (with a validation RMSE of 0.22 eV) performs better than scaling relations. The best descriptor among the primary features (the SISSO 1D, Φ_0 descriptor) is found to be the d-band center, i.e. SISSO also identifies the physics that has already been discovered in form of the d-band model more than 20 years ago.[26]

A comparison of the performance of the d-band center, scaling relations, and the best $(8D, \Phi_3)$ SISSO descriptor on the test data set (the remaining 50% of the SA and AB alloys data) is shown in Figure 3(b). The RMSEs are d-band center: 0.37 eV, scaling relations: 0.28 eV, SISSO: 0.15 eV, and the maximum absolute errors (maxAEs) are d-band center: 1.31 eV, scaling relations: 1.43 eV, SISSO: 0.61 eV. Here, the maxAE is the maximal error observed for any of the adsorbates considered in the multitask learning. As already observed for the validation data, the test results thus demonstrate the great improvement of the new SISSO descriptor compared to previous approaches.

We mention here also the performance of scaling relations when considering only the most stable adsorption site of each adsorbate at each metal or alloy as this is often perceived to be most relevant for catalysis (see Supplementary Figure S5). The RMSE on the pure metals (alloys) is 0.16 eV (0.17 eV). As already mentioned, this scaling-relation-based approach is very expensive because, to screen an unknown alloy, all potential adsorption sites must be tested to find the most stable adsorption sites of the descriptors. For comparison, the predictions from the best $(8D, \Phi_3)$ SISSO descriptor on the pure metals in the training data set (alloys in the test data set) is 0.09 eV (0.15 eV), i.e. even for this restricted case the SISSO approach is more accurate in addition to being much less costly.

In Supplementary Table S7, we provide additional information about the largest deviations between calculated and SISSO-predicted adsorption energies for the alloys. In general, the C adsorption energies are the most difficult to predict because they vary much more over the TM series than, e.g., the H adsorption energies. In addition, the most difficult alloys to predict adsorption energies for are those that combine a more noble and a more reactive metal such as Au@Ni or Ag@Cu. The maximum absolute error of 0.61 eV is found for the OH adsorption energy on top of the Au atom in the Au@ Ni alloy. The embedding of the larger and more noble Au atom in the smaller lattice constant and more reactive Ni surface (see Figure 1(b)) probably provides an adsorption site that is both geometrically and electronically very different from everything else in the data set and thus harder to predict.

Transferability of Descriptors. To further test the predictive performance of the best $(8D, \Phi_3)$ SISSO descriptor, we show in Figure 3(c) the error distribution for the prediction of adsorption energies on three new facets, the (111), (110),

and (100) facets. Some sites, in particular those found on the (111) facet, are very similar to sites found on the (211) facet in the training data. However, the (110) and (100) facets contain sites that are very different, albeit of similar coordination numbers (7−9.5), compared to the (211) facet sites. It is seen that the descriptor with the best performance for (211) facets performs very well for the (111) facet (RMSE of 0.17 eV) but significantly worse for the (110) and (100) facets (RMSEs of 0.29 and 0.30 eV, respectively). This shows a well-known limitation of compressed sensing (and machine learning) methodologies, namely that a good transferability cannot be expected *a priori* for cases not previously encountered in the training or validation data. However, to put this in perspective, we note that this "poor" RMSE for the other facets is essentially of the same level as the RMSEs obtained for the widely used scaling relations in the first place.
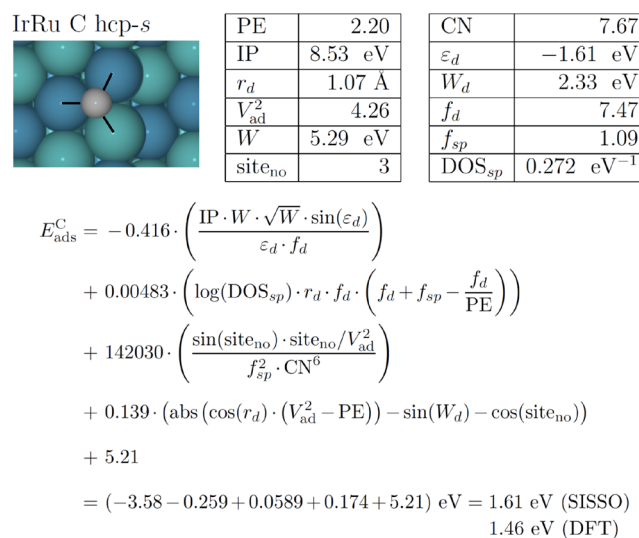
To identify a descriptor that has a good compromise between accuracy for alloys and facets, we now include 50% of the facets data in the validation data (see Supplementary Figure S6). The new best descriptor is found for the hyper parameters 8D, Φ_1. It is interesting to note that because this is a Φ_1 descriptor, the functional form of the features is much less complex than for the Φ_3 descriptor optimized for the alloys alone. This suggests that a less complex mathematical form is required for a descriptor that is transferable across both alloys and active site motifs.

The RMSEs (maxAEs) of the (8D, Φ_1) descriptor for the test set (the remaining 50% of the facets data) are found to be (111): 0.17 eV (0.44 eV), (110): 0.21 eV (0.59 eV), (100): 0.24 eV (0.48 eV). These very moderate errors show that it is possible to identify a descriptor with a good predictive performance for a wide range of structural motifs as exemplified by the low-index fcc facets. The improved performance on the facets data sets comes at the very moderate expense of increasing the RMSEs for the alloys to 0.18 eV compared to 0.15 eV before. We therefore suggest the (8D, Φ_1) SISSO descriptor for cases where simultaneous screening of alloys and a wide range of active site motifs is desired. It should be emphasized though that in general a good performance can only be expected for active site types that resemble to some extent those types for which the descriptor was optimized. Likely, more varied training and validation data will be required to identify a descriptor that would work for very different active site motifs such as kinks, vacancies, and adatoms.

Composition of Descriptors. Having confirmed the predictive performance of the identified descriptors, we now move on to discuss their composition. As an example of a descriptor identified by SISSO, we give in eq 1 the best 4D descriptor of Φ_3 from Figure 3(a) (alloy validation data set):

$$P^j = c_1^j \cdot \left(\frac{IP \cdot W \cdot \sqrt{W} \cdot \sin(\varepsilon_d)}{\varepsilon_d \cdot f_d} \right)$$
$$+ c_2^j \cdot \left(\log(DOS_{sp}) \cdot r_d \cdot f_d \cdot \left(f_d + f_{sp} - \frac{f_d}{PE} \right) \right)$$
$$+ c_3^j \cdot \left(\frac{\sin(\text{site}_{no}) \cdot \text{site}_{no}/V_{ad}^2}{f_{sp}^2 \cdot CN^6} \right)$$
$$+ c_4^j \cdot \left(abs\left(\cos(r_d) \cdot \left(V_{ad}^2 - PE \right) \right) - \sin(W_d) - \cos(\text{site}_{no}) \right)$$
$$+ c_0^j \quad , \tag{1}$$

where P^j is an adsorption energy of adsorbate j and the primary features entering the descriptor are evaluated for the material/ site combination relevant for P^j. The d-band center and sp-band properties such as the Pauling electronegativity are identified as highly important primary features. This was also found in previous studies employing artificial neural networks,[6,7] but it is here expressed in an explicit nonlinear functional form owing to the compressed sensing methodology. Among the remaining primary features entering the descriptor, we especially highlight the DOS at the Fermi level (here of the sp-band), which is a feature not considered in these previous studies, even though its importance for the reactivity of TM surfaces was discussed already more than 30 years ago by Yang and Parr.[25] In Figure 4, we provide an example adsorption energy prediction using eq 1.

IrRu C hcp-*s*

PE	2.20		CN	7.67	
IP	8.53 eV		ε_d	−1.61 eV	
r_d	1.07 Å		W_d	2.33 eV	
V_{ad}^2	4.26		f_d	7.47	
W	5.29 eV		f_{sp}	1.09	
site_{no}	3		DOS_{sp}	0.272 eV^{-1}	

$$E_{ads}^C = -0.416 \cdot \left(\frac{IP \cdot W \cdot \sqrt{W} \cdot \sin(\varepsilon_d)}{\varepsilon_d \cdot f_d} \right)$$
$$+ 0.00483 \cdot \left(\log(DOS_{sp}) \cdot r_d \cdot f_d \cdot \left(f_d + f_{sp} - \frac{f_d}{PE} \right) \right)$$
$$+ 142030 \cdot \left(\frac{\sin(\text{site}_{no}) \cdot \text{site}_{no}/V_{ad}^2}{f_{sp}^2 \cdot CN^6} \right)$$
$$+ 0.139 \cdot \left(abs\left(\cos(r_d) \cdot \left(V_{ad}^2 - PE \right) \right) - \sin(W_d) - \cos(\text{site}_{no}) \right)$$
$$+ 5.21$$
$$= (-3.58 - 0.259 + 0.0589 + 0.174 + 5.21) \text{ eV} = 1.61 \text{ eV (SISSO)}$$
$$1.46 \text{ eV (DFT)}$$

Figure 4. Example of SISSO adsorption energy prediction for C at an hcp-s site of the IrRu alloy using the 4D descriptor of Φ_3 given in eq 1. The tabulated primary features are calculated as averages over the three metal atoms (two Ir atoms and one Ru atom) making up the IrRu hcp-s site (marked with black lines). The shown fitting coefficients c_i^C are specific for C. For ease of reading, their units are not shown; these depend on the units of the primary features entering each feature to ensure that the adsorption energy comes out in eV.

An overview of the identified descriptors of each dimension and rung for both the alloy validation data set and the combined alloy and facets validation data set together with the fitting coefficients c_i^j is given in Supplementary Tables S8 and S9.

Enlarging the Training Data Set. The predictive performance of the identified descriptors for alloy screening is already impressive, given that no explicit information on alloys was given in the training data. However, a further advantage of data-driven approaches is that the learning can be systematically improved by enlarging the training data set. In contrast, the rigid format of linear scaling relations does not allow for significant improvements, even if fitting also to the alloys data, as evidenced by the scattering of the alloy data points around the fitted line in Figure 2(a) and (b).

To provide a simple estimate of the learning improvement possible when also including alloys in the training data, we identify a new SISSO descriptor (see Supplementary Table S10) using as training data, specifically the pooled metal and

alloys data sets, but excluding the 23 DFT-calculated alloy data points (colored stars) shown in Figures 2(a) and (b). For this, we use the hyperparameters that were found to be best for alloys (8D, Φ_3). The colored dots in Figures 2(a) and (b) show the SISSO predictions for the test set (the data points left out in the training). Already a visual inspection reveals that the agreement is very good. The maxAE (of 0.52 eV) is found for the OH adsorption energy of the dark green point in Figure 2(a), which corresponds to OH adsorption on top of the Au atom in the Au@Ni alloy. Note that exactly this data point is also the maxAE (with the slightly larger value of 0.61 eV) for the descriptor trained only on the pure metals. The RMSE over the 23 predicted alloy data points in Figures 2(a) and (b) decreases from 0.23 eV (training on pure metals only) to 0.18 eV (training on pooled data set).

Overall, the good agreement between the DFT-calculated values and the SISSO predictions shows that our models have the required accuracy to systematically search for outliers to scaling relations. In addition, our approach is computationally cheap enough to allow for the screening of immense alloy spaces. For this, we will next present a simple example.

A First Screening Example. In the following, we make use of a SISSO descriptor (see Supplementary Table S11) identified using the hyperparameters (8D, Φ_3) and the entire pooled metals and alloys data sets for training. We predict the adsorption energies for the adsorbates and sites considered in Figures 2(a) and (b) on the additional 32 possible AB alloys (those that were not explicitly calculated by DFT) as shown with the gray dots. Similar to the explicitly DFT-calculated alloy data points, there is a considerable scatter around the scaling relation lines. This shows that there indeed exist many materials with potentially interesting catalytic properties which would be missed by a scaling-relation-based screening approach.

A particularly interesting perspective for catalyst screening is to be able to search directly for candidate materials that break scaling relations. For many reactions, the incentive would be to break scaling relations in a desired way because it has been suggested that scaling relations impose an upper limit to the possible catalyst activity. For example, it is known that for the oxygen evolution reaction, it would be desirable to find a material where O is destabilized relative to OOH,[38] and for electrochemical CO_2 reduction, it is desirable to destabilize CO relative to CHO.[39] For other reactions, where optimum catalytic activity has hitherto been exclusively formulated in terms of singular descriptors, the interest would be to evaluate the effect of scatter in the binding of other important intermediates that hitherto has been assumed as fixed through scaling relations. For the oxygen reduction reaction (ORR), optimum catalytic activity has for instance been associated with an optimum oxygen adsorption energy,[40] while for selective ethanol synthesis, both the carbon and the oxygen adsorption energy must be simultaneously optimized.[37] The activity at the top of this theoretical volcano curve is then independent of e.g. the OH (ORR) or CH (ethanol) adsorption energy, as the latter are connected to the optimum O or C adsorption energy through a scaling relation, respectively.

In Figures 2(c) and (d) we specifically check on the scatter by showing histograms of the predicted adsorption energies for (c) O and OH and (d) C and CH at all potential adsorption sites of all 36 AB alloys. There are more predicted points for OH (around 1000) than for the other adsorbates (around 400) because OH adsorption can take place at more site types, i.e.

also at top and bridge sites. The black shaded areas highlighted in Figure 2(c) the ORR "volcano optimal" O adsorption energy on (111) facets[36] and in Figure 2(d) the optimal C adsorption energy on (211) facets for selective ethanol synthesis.[37] For this simple screening example, we assume that only the most stable adsorption site of a given adsorbate plays a catalytic role, keeping in mind that in reality, less stable (metastable) sites also could get populated at higher coverages. The SISSO approach directly gives us the energetics for the most stable and all metastable sites, so in general, we are not limited to considering only most stable sites. Because the ORR volcano was developed for (111) facets, we search for materials for which the most stable O adsorption energy among the (111)-like (terrace) sites of the (211) facet falls within the desired range. This results in three candidate materials: PdPt, AgPd, and AgPt. The latter material is highlighted in Figure 2(c) because it has an OH adsorption energy on its most stable bridge1-t site that is 0.23 eV lower than the value that would be predicted for this site from scaling relations (indicated by the black arrow) based on the O adsorption energy on its most stable hcp-t site. The opposite behavior (a lower O adsorption energy relative to OH) is seen for the material shown with the green dot (PdPt). The cause of this breaking of scaling relations is thereby the slightly different oxygenate adsorption energy for Pt and Ag, and the fact that the preferred adsorption sites for OH and O have a different composition of Pt and Ag. It is important to note that our approach does not assume that either O or OH adsorption are appropriate descriptors; instead, we include both of them as independent options. A similar breaking of scaling relations is observed for the ethanol synthesis example. Here, SISSO recovers the scatter in the most stable CH adsorption energies for materials (RuAg, RhAu, RuAu, IrAu, and CuIr) that all have about the same most stable "optimum" C adsorption energy. This scatter shows the extent to which it is possible to tune the CH adsorption energy independently of the C adsorption energy and thereby further tailor the catalytic activity.

A high-throughput screening perspective. It should be noted that the moderate breaking of scaling relations observed in the simple examples from the previous section is related to the consideration of only a handful of "near-optimal" materials (out of a total of only 36 considered materials) and in particular the consideration of only most stable adsorption sites. A full assessment of the extent to which scaling relations can be broken on alloys with mixed-metal surfaces would be revealed only by a full high-throughput screening of hundreds of thousands of materials that is beyond the scope of the present study. Such a high-throughput screening could also involve the evaluation of a microkinetic model for each catalyst material that takes into account all possible adsorption sites of every adsorbate as well as kinetic barriers for reaction and diffusion steps through BEP relations. If such a microkinetic model was initially carried out within the simplifying mean-field approximation,[41] the cost of its evaluation would still only be a negligible fraction of the (already small) cost of carrying out a DFT calculation of the primary features of the clean catalyst surface for the descriptor evaluation. Once a selection of promising catalyst materials had been identified, a next step could then be the evaluation of a more thorough microkinetic model from, e.g., a kinetic Monte Carlo simulation,[32] possibly also taking into account lateral interactions between the adsorbates through cluster expansion methods.[42] A full assessment of identified promising catalyst materials would

ultimately also need to take into account other aspects such as bulk and surface segregation stability under realistic surface coverages for the chemical reaction and reaction conditions of interest as well as for instance stability against metal stripping in electrocatalysis applications.

CONCLUSIONS

In summary, we used compressed sensing to identify new and better descriptors that allow prediction of adsorption energies for a whole range of atoms and molecules at all potential surface sites of TMs and bimetallics formed of TMs. The descriptors can be obtained from a single DFT calculation of the clean surface, and their predictive power extends over both multimetallics and various surface facets. Importantly, this enables low-cost catalyst screening in not only materials but also active site space[43] with unprecedented accuracy. With respect to materials, the thereby enabled systematic identification and analysis of outliers to traditional scaling relation energetics seems particularly promising. With respect to active sites, the availability of energetic data for a wide range of site types paves the way to actively embrace the uncertainty in surface structure and composition of working catalysts.[3]

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.8b04478.

> Additional DFT and SISSO computational details, multitask learning versus learning of separate descriptors, overview of largest prediction errors, all scaling relation plots, tested SIS values, hyperparameter testing with facets data in validation data set, and all identified descriptors (ZIP)

AUTHOR INFORMATION

Corresponding Author
*E-mail: mie.andersen@ch.tum.de.
ORCID Ⓘ
Mie Andersen: 0000-0002-9943-1534
Karsten Reuter: 0000-0001-8473-8659
Present Address
⊥Center for Energy Science and Technology, Skolkovo Institute of Science and Technology, 143026 Moscow, Russia.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Ertl, G. Reactions at Surfaces: From Atoms to Complexity (Nobel Lecture). *Angew. Chem., Int. Ed.* **2008**, *47*, 3524−3535.

(2) Reuter, K.; Frenkel, D.; Scheffler, M. The Steady State of Heterogeneous Catalysis, Studied by First-Principles Statistical Mechanics. *Phys. Rev. Lett.* **2004**, *93*, 116105.

(3) Reuter, K. Ab Initio Thermodynamics and First-Principles Microkinetics for Surface Catalysis. *Catal. Lett.* **2016**, *146*, 541−563.

(4) Nørskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. Towards the Computational Design of Solid Catalysts. *Nat. Chem.* **2009**, *1*, 37−46.

(5) Draxl, C.; Scheffler, M. NOMAD: The FAIR concept for big data-driven materials science. *MRS Bull.* **2018**, *43*, 676−682.

(6) Ma, X.; Li, Z.; Achenie, L. E. K.; Xin, H. Machine-Learning-Augmented Chemisorption Model for CO2 Electroreduction Catalyst Screening. *J. Phys. Chem. Lett.* **2015**, *6*, 3528−3533.

(7) Li, Z.; Wang, S.; Chin, W. S.; Achenie, L. E.; Xin, H. High-Throughput Screening of Bimetallic Catalysts Enabled by Machine Learning. *J. Mater. Chem. A* **2017**, *5*, 24131−24138.

(8) Ulissi, Z. W.; Tang, M. T.; Xiao, J.; Liu, X.; Torelli, D. A.; Karamad, M.; Cummins, K.; Hahn, C.; Lewis, N. S.; Jaramillo, T. F.; Chan, K.; Nørskov, J. K. Machine-Learning Methods Enable Exhaustive Searches for Active Bimetallic Facets and Reveal Active Site Motifs for CO2 Reduction. *ACS Catal.* **2017**, *7*, 6600−6608.

(9) Gasper, R.; Shi, H.; Ramasubramaniam, A. Adsorption of CO on Low-Energy, Low-Symmetry Pt Nanoparticles: Energy Decomposition Analysis and Prediction via Machine-Learning Models. *J. Phys. Chem. C* **2017**, *121*, 5612−5619.

(10) Jinnouchi, R.; Asahi, R. Predicting Catalytic Activity of Nanoparticles by a DFT-Aided Machine-Learning Algorithm. *J. Phys. Chem. Lett.* **2017**, *8*, 4279−4283.

(11) Tran, K.; Ulissi, Z. W. Active Learning Across Intermetallics to Guide Discovery of Electrocatalysts for CO2 Reduction and H2 Evolution. *Nat. Catal.* **2018**, *1*, 696−703.

(12) 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** DOI: 10.1016/j.joule.2018.12.015.

(13) Abild-Pedersen, F.; Greeley, J.; Studt, F.; Rossmeisl, J.; Munter, T. R.; Moses, P. G.; Skúlason, E.; Bligaard, T.; Nørskov, J. K. Scaling Properties of Adsorption Energies for Hydrogen-Containing Molecules on Transition-Metal Surfaces. *Phys. Rev. Lett.* **2007**, *99*, No. 016105.

(14) Grabow, L. C. When Outliers Make All The Difference. *ChemCatChem* **2012**, *4*, 1887−1888.

(15) Kyriakou, G.; Boucher, M. B.; Jewell, A. D.; Lewis, E. A.; Lawton, T. J.; Baber, A. E.; Tierney, H. L.; Flytzani-Stephanopoulos, M.; Sykes, E. C. H. Isolated Metal Atom Geometries as a Strategy for Selective Heterogeneous Hydrogenations. *Science* **2012**, *335*, 1209−1212.

(16) Darby, M. T.; Réocreux, R.; Sykes, E. C. H.; Michaelides, A.; Stamatakis, M. Elucidating the Stability and Reactivity of Surface Intermediates on Single-Atom Alloy Catalysts. *ACS Catal.* **2018**, *8*, 5038−5050.

(17) Ouyang, R.; Curtarolo, S.; Ahmetcik, E.; Scheffler, M.; Ghiringhelli, L. M. SISSO: A Compressed-sensing Method for Identifying the Best Low-Dimensional Descriptor in an Immensity of Offered Candidates. *Phys. Rev. Mater.* **2018**, *2*, No. 083802.

(18) Ouyang, R.; Ahmetcik, E.; Carbogno, C.; Scheffler, M.; Ghiringhelli, L. M. Simultaneous Learning of Several Materials Properties from Incomplete Databases with Multi-Task SISSO. *ArXiv e-prints 1901.00948* **2019** DOI: 10.1088/2515-7639/ab077b.

(19) Giannozzi, P.; et al. QUANTUM ESPRESSO: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.

(20) Wellendorff, J.; Lundgaard, K. T.; Møgelhøj, A.; Petzold, V.; Landis, D. D.; Nørskov, J. K.; Bligaard, T.; Jacobsen, K. W. Density Functionals for Surface Science: Exchange-Correlation Model Development with Bayesian Error Estimation. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2012**, *85*, 235149.

(21) Han, J. W.; Kitchin, J. R.; Sholl, D. S. Step Decoration of Chiral Metal Surfaces. *J. Chem. Phys.* **2009**, *130*, 124710.

(22) Curtarolo, S.; Setyawan, W.; Wang, S.; Xue, J.; Yang, K.; Taylor, R. H.; Nelson, L. J.; Hart, G. L.; Sanvito, S.; Buongiorno-Nardelli, M.; Mingo, N.; Levy, O. AFLOWLIB.ORG: A Distributed Materials Properties Repository from High-Throughput Ab Initio Calculations. *Comput. Mater. Sci.* **2012**, *58*, 227−235.

(23) The complete DFT data set can also be freely downloaded from the NOMAD Repository [http://dx.doi.org/10.17172/NOMAD/2019.02.22-1].

(24) Fukui, K. Role of Frontier Orbitals in Chemical Reactions. *Science* **1982**, *218*, 747−754.

(25) Yang, W.; Parr, R. G. Hardness, Softness, and the Fukui Function in the Electronic Theory of Metals and Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 6723−6726.

(26) Hammer, B.; Nørskov, J. K. Electronic Factors Determining the Reactivity of Metal Surfaces. *Surf. Sci.* **1995**, *343*, 211.

(27) Ruban, A.; Hammer, B.; Stoltze, P.; Skriver, H.; Nørskov, J. Surface Electronic Structure and Reactivity of Transition and Noble Metals. *J. Mol. Catal. A: Chem.* **1997**, *115*, 421−429.

(28) Xin, H.; Holewinski, A.; Linic, S. Predictive Structure−Reactivity Models for Rapid Screening of Pt-Based Multimetallic Electrocatalysts for the Oxygen Reduction Reaction. *ACS Catal.* **2012**, *2*, 12−16.

(29) Xin, H.; Vojvodic, A.; Voss, J.; Nørskov, J. K.; Abild-Pedersen, F. Effects of d-Band Shape on the Surface Reactivity of Transition-Metal Alloys. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *89*, 115114.

(30) Calle-Vallejo, F.; Loffreda, D.; Koper, M. T. M.; Sautet, P. Introducing Structural Sensitivity into Adsorption-Energy Scaling Relations by Means of Coordination Numbers. *Nat. Chem.* **2015**, *7*, 403−410.

(31) Roling, L. T.; Abild-Pedersen, F. Structure-Sensitive Scaling Relations: Adsorption Energies from Surface Site Stability. *ChemCatChem* **2018**, *10*, 1643−1650.

(32) Andersen, M.; Plaisance, C. P.; Reuter, K. Assessment of Mean-Field Microkinetic Models for CO Methanation on Stepped Metal Surfaces Using Accelerated Kinetic Monte Carlo. *J. Chem. Phys.* **2017**, *147*, 152705.

(33) Studt, F.; Abild-Pedersen, F.; Wu, Q.; Jensen, A. D.; Temel, B.; Grunwaldt, J. D.; Nørskov, J. K. CO Hydrogenation to Methanol on Cu−Ni Catalysts: Theory and Experiment. *J. Catal.* **2012**, *293*, 51−60.

(34) Calle-Vallejo, F.; Tymoczko, J.; Colic, V.; Vu, Q. H.; Pohl, M. D.; Morgenstern, K.; Loffreda, D.; Sautet, P.; Schuhmann, W.; Bandarenka, A. S. Finding Optimal Surface Sites on Heterogeneous Catalysts by Counting Nearest Neighbors. *Science* **2015**, *350*, 185−189.

(35) Jørgensen, M.; Grönbeck, H. The Site-Assembly Determines Catalytic Activity of Nanoparticles. *Angew. Chem., Int. Ed.* **2018**, *57*, 5086−5089.

(36) Nørskov, J. K.; Rossmeisl, J.; Logadottir, A.; Lindqvist, L.; Kitchin, J. R.; Bligaard, T.; Jonsson, H. Origin of the Overpotential for Oxygen Reduction at a Fuel-Cell Cathode. *J. Phys. Chem. B* **2004**, *108*, 17886−17892.

(37) Medford, A. J.; Lausche, A. C.; Abild-Pedersen, F.; Temel, B.; Schjødt, N. C.; Nørskov, J. K.; Studt, F. Activity and Selectivity Trends in Synthesis Gas Conversion to Higher Alcohols. *Top. Catal.* **2014**, *57*, 135−142.

(38) Rossmeisl, J.; Logadottir, A.; Nørskov, J. Electrolysis of Water on (Oxidized) Metal Surfaces. *Chem. Phys.* **2005**, *319*, 178−184.

(39) Li, Y.; Sun, Q. Recent Advances in Breaking Scaling Relations for Effective Electrochemical Conversion of CO2. *Adv. Energy Mater.* **2016**, *6*, 1600463.

(40) Greeley, J.; Stephens, I. E. L.; Bondarenko, A. S.; Johansson, T. P.; Hansen, H. A.; Jaramillo, T. F.; Rossmeisl, J.; Chorkendorff, I.; Nørskov, J. K. Alloys of Platinum and Early Transition Metals as Oxygen Reduction Electrocatalysts. *Nat. Chem.* **2009**, *1*, 552−556.

(41) Medford, A. J.; Shi, C.; Hoffmann, M. J.; Lausche, A. C.; Fitzgibbon, S. R.; Bligaard, T.; Nørskov, J. K. CatMAP: A Software Package for Descriptor-Based Microkinetic Mapping of Catalytic Trends. *Catal. Lett.* **2015**, *145*, 794−807.

(42) Stamatakis, M.; Piccinin, S. Rationalizing the Relation between Adlayer Structure and Observed Kinetics in Catalysis. *ACS Catal.* **2016**, *6*, 2105−2111.

(43) Reuter, K.; Plaisance, C. P.; Oberhofer, H.; Andersen, M. Perspective: On the Active Site Model in Computational Catalyst Screening. *J. Chem. Phys.* **2017**, *146*, No. 040901.