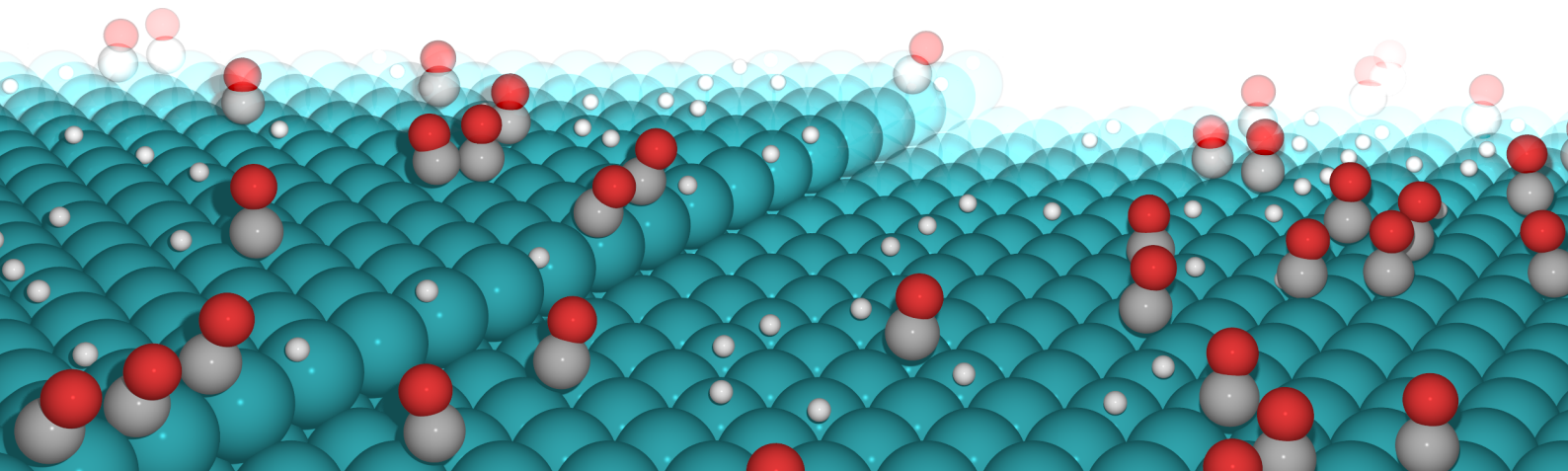
Technische Universität München
TUM School of Natural Sciences

Ab Initio Studies of the Activity and Selectivity of Transition Metal Catalysts for CO Hydrogenation

Martin Maximilian Deimel

Dissertation



Technische Universität München
TUM School of Natural Sciences

Ab Initio Studies of the Activity and Selectivity of Transition Metal Catalysts for CO Hydrogenation

Martin Maximilian Deimel

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Christopher Stein

Prüfer*innen der Dissertation:

1. Prof. Dr. Karsten Reuter
2. Assoc. Prof. Dr. Mie Andersen

Die Dissertation wurde am 29.11.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 04.01.2023 angenommen.

„Aber ich finde trotzdem immer Trost in Wissenschaft.
Mit jeder Frage, die man beantwortet, tun sich mindestens drei neue auf.
Man erklärt sich die Welt nicht weg, sondern ich finde, man macht sie immer wundervoller.“[1]

– Mai Thi Nguyen-Kim –

Preface

This thesis is publication-based, which means that its content provides both a framework around as well as a summary of two peer-reviewed articles published in international scientific journals. The original articles titled *"Active Site Representation in First-Principles Microkinetic Models: Data-Enhanced Computational Screening for Improved Methanation Catalysts"*[2] and *"Selectivity Trends and Role of Adsorbate–Adsorbate Interactions in CO Hydrogenation on Rhodium Catalysts"*[3] are both published in *ACS Catalysis* and appended to this thesis. They are accompanied by an introduction to the theoretical basics and corresponding relevant literature.

The work presented was performed between March 2018 and November 2022 at the Chair of Theoretical Chemistry and Catalysis Research Center of the Technical University of Munich, and the Theory Department of the Fritz Haber Institute of the Max Planck Society under the supervision of Assoc. Prof. Dr. Mie Andersen and Prof. Dr. Karsten Reuter.

Munich, November 2022

Abstract

First-principles (1p)-based computational modeling of surface chemical reactions has matured into a predictive-quality instrument for the study of heterogeneous catalysis. While empirical observations of different catalyst materials could be successfully reconciled with theoretical calculations on the macroscopic scale, the intricate processes and interactions on the microscopic scale still pose a challenging field of research. Although important milestones on the way to an in-depth understanding of a catalytic process have been accomplished, the growing availability of computational resources and the ongoing development of algorithms and data-driven methods pave the way for the efficient computational modeling of highly complex microkinetic reaction networks.

Based on the energetic input from density functional theory (DFT) and various scaling relations, an efficient, yet approximate method for the modeling of catalytic reactions is the mean field approximation (MFA). In combination with a reductionist representation of the catalytic surface under investigation, it served as a reliable and successful tool for the screening and evaluation of activity trends of catalyst materials while keeping the computational cost tractable. This cumulative thesis describes the basics of such an MFA-based approach unveiling its shortcomings and the concomitant need for a more detailed active site representation. Additionally, a data-driven approach to predict the adsorption energetics is described, overcoming the inherent shortcomings of the linear scaling relations when applied to mixed-metal catalysts. Here, a comparison of models with different levels of active site representation illustrates the importance of the active site resolution on binary alloy catalysts and allows for the identification of possible improved methanation catalysts due to synergistic effects, which cannot be described in a more coarse-grained description.

An MFA-based approach avoids an explicit spatially resolved surface representation and the concomitant individual interactions between adsorbates, and consequently precludes an in-depth understanding of the microscopic mechanisms on the catalyst surface. While such an approach has successfully aided in the prediction of activities and activity trends among different catalysts facilitating large-scale screening studies, it is not capable of accurately including adsorbate interactions and often fails to explain selectivity trends. To capture these, a more detailed description on the microscale is required. This is realized by a kinetic Monte Carlo (KMC) approach, which properly accounts for the probabilistic character of microkinetics by simulating the underlying Markov process. For the catalytic activity and selectivity of the carbon monoxide (CO) hydrogenation reaction on Rh catalysts, a surprising dependency of the activity and selectivity on correlations among the adsorbed species influencing the coverage is found, caused by subtle variations in adsorption energetics originating from attractive or repulsive interactions. This demonstrates the essential importance of accurately describing interactions and local coverages on the atomic scale, as a catalyst usually performs best in regions of intermediate coverages.

Zusammenfassung

Die auf ersten Prinzipien (1p) basierende computergestützte Modellierung chemischer Oberflächenreaktionen hat sich zu einem vorhersagekräftigen Werkzeug für die Untersuchung der heterogenen Katalyse entwickelt. Während empirische Beobachtungen unterschiedlicher Katalysatormaterialien erfolgreich auf makroskopischer Ebene mit theoretischen Berechnungen in Einklang gebracht werden konnten, stellen auf mikroskopischer Ebene die komplexen Prozesse und Wechselwirkungen nach wie vor ein herausforderndes Forschungsgebiet dar. Obwohl bereits wichtige Etappenziele zu einem tiefgreifenden Verständnis eines katalytischen Prozesses erreicht wurden, ebnen die zunehmende Verfügbarkeit von Rechenressourcen sowie die kontinuierliche Entwicklung von Algorithmen und datengestützten Methoden den Weg für die effiziente computergestützte Modellierung hochkomplexer mikrokinetischer Reaktionsnetzwerke.

Basierend auf dem energetischen Input aus der Dichtefunktionaltheorie (DFT) und verschiedenen Skalierungszusammenhängen, stellt die Mean-Field-Approximation (MFA) eine effiziente, aber dennoch approximative Methode für die Modellierung katalytischer Reaktionen dar. Sie diente in Kombination mit einer reduktionistischen Darstellung der zu untersuchenden katalytischen Oberfläche als zuverlässiges und erfolgreiches Werkzeug für das Screening und die Evaluierung von Aktivitätstrends von Katalysatormaterialien bei gleichzeitig tragbarem Rechenaufwand. In dieser kumulativen Dissertation werden die Grundlagen eines solchen MFA-basierten Ansatzes beschrieben sowie seine Schwächen und die damit einhergehende Notwendigkeit einer detaillierteren Darstellung der aktiven Zentren aufgezeigt. Darüber hinaus wird ein datengestützter Ansatz zur Vorhersage der Adsorptionsenergetik beschrieben, der die inhärenten Schwächen der linearen Skalierungszusammenhänge bei der Anwendung auf Katalysatoren, welche aus verschiedenen Metallen bestehen, überwindet. Ein Vergleich von Modellen mit unterschiedlicher Darstellungsgenauigkeit der aktiven Zentren zeigt die Wichtigkeit dieser Darstellungsgenauigkeit bei Katalysatoren aus binären Legierungen auf. Dieser Vergleich ermöglicht die Ermittlung leistungsfähigerer Katalysatoren für die Methanisierung aufgrund von synergistischen Effekten, die mit einer geringeren Darstellungsgenauigkeit nicht beschrieben werden können.

Ein auf der MFA basierender Ansatz vermeidet eine explizite ortsaufgelöste Oberflächendarstellung und die damit einhergehenden spezifischen Wechselwirkungen zwischen Adsorbaten. Folglich verhindert sie tiefgreifendere Erkenntnisse über die mikroskopischen Mechanismen auf der Katalysatoroberfläche. Obwohl ein solcher Ansatz erfolgreich zur Vorhersage von Aktivitäten und Aktivitätstrends zwischen verschiedenen Katalysatoren beigetragen und umfangreiche Screenings ermöglicht hat, kann er Wechselwirkungen zwischen Adsorbaten nicht genau berücksichtigen und scheitert oft daran Selektivitätstrends zu erklären. Um diese Wechselwirkungen zu erfassen, ist eine detailliertere Darstellung auf mikroskopischer Ebene erforderlich. Dies wird durch einen kinetischen Monte-Carlo (KMC) Ansatz erreicht, der durch Simulation des zugrunde liegenden Markov-Prozesses den probabilistischen Charakter der Mikrokinetik korrekt abbildet. Für die katalytische Aktivität und Selektivität der Kohlenmonoxid (CO)-Hydrierungsreaktion auf Rh-Katalysatoren wird eine

überraschende Abhängigkeit der Aktivität und Selektivität von Korrelationen zwischen den adsorbierten Spezies ermittelt, die die Bedeckung beeinflussen und durch feine Änderungen der Adsorptionsenergien verursacht werden, die auf attraktive oder abstoßende Wechselwirkungen zurückzuführen sind. Dies demonstriert, wie wichtig es ist, Wechselwirkungen und lokale Bedeckungen auf atomarer Ebene genau zu beschreiben, da ein Katalysator für gewöhnlich in Bereichen mittlerer Bedeckung das beste Leistungspotenzial entfaltet.

Contents

List of Figures

List of Tables

List of Abbreviations

1p	first-principles
AFLOW	Automatic Flow
AI	artificial intelligence
C_nH_{2n+2}	alkane
NH_3	ammonia
bcc	body-centered cubic
BEEF-vdW	Bayesian error estimation functional with van der Waals correlation
BEP	Brønsted-Evans-Polanyi
CO_2	carbon dioxide
CO	carbon monoxide
CatMAP	Catalysis Microkinetic Analysis Package
CE	cluster expansion
CHULL	convex hull
CI	climbing image
CMR	Computational Materials Repository
CS	compressed sensing
DFT	density functional theory
DOF	degree of freedom
DRC	degree of rate control
fcc	face-centered cubic
FS	final state
GGA	generalized gradient approximation
hcp	hexagonal close packed
H_2	hydrogen
IS	initial state
KMC	kinetic Monte Carlo
LDA	local density approximation
MAE	mean absolute error
MEP	minimum energy path
CH_4	methane
MFA	mean field approximation
MKM	microkinetic modeling
ML	machine learning
NEB	nudged elastic band
N_2	nitrogen

NOMAD	Novel Materials Discovery
OQMD	Open Quantum Materials Database
otf	on-the-fly
O_2	oxygen
PES	potential energy surface
RDS	rate-determining step
RLS	rate-limiting step
RMSE	root mean squared error
SISSO	Sure Independence Screening and Sparsifying Operator
syngas	synthesis gas
TM	transition metal
TOF	turnover frequency
TST	transition state theory
TS	transition state
VSSM	variable step-size method
H_2O	water
XPK	extended phenomenological kinetics
ZPE	zero-point energy

1 Introduction

In 1823, Johann Wolfgang Döbereiner discovered the ignition of hydrogen (H_2) gas when in contact with platinum powder and described his findings as a "highly curious property" emphasizing the lack of knowledge about catalytic processes just two centuries ago.[4, 5] Since then, catalysis research and catalytic processes emerged to be an essential part of our society and economy. Not only do we owe economic and ecologic advances in technology to catalytic processes, but also the possibility to feed an ever growing population.[5] As one intriguing success story, the later development was mainly made possible by the research of Fritz Haber, Carl Bosch, and Alwin Mittasch to synthesize ammonia (NH_3) from atmospheric nitrogen (N_2) and H_2 on an Fe catalyst.[5, 6] Another important process from the energy sector is the Fischer-Tropsch synthesis to produce hydrocarbons from synthesis gas (syngas) – a mixture of carbon monoxide (CO) and H_2.[7] However, fossil energy carriers, such as coal or natural gas, serve as the main source materials for the production of syngas. In order to shift from fossil feedstock to renewable energy carriers, such source materials can also be produced more environmentally friendly from biological feed stock.[8] Additionally, catalyst materials often consist of rare and/or problematic components. For that reason, an important task is to develop improved catalysts with low ecologic and economic adversarial impact.[9, 10]

Before the development of computational methods, catalyst development mainly comprised of a trial and error approach of hypothetically assumed candidates in a laboratory setup.[5, 11] As the catalyst does not change the thermodynamics of a reaction but the kinetics, methods and theories have been developed to qualitatively and quantitatively specify the effect.[12] With the advent of *in silico* methods and their subsequent improvement, together with the increasing availability of computational resources, the field of computational chemistry has turned into an accurate and predictive tool for the investigation of the properties and characteristics of catalysts, supporting increasing efforts for rational catalyst design.[13–15] One key aspect of computational modeling is the method on how to evaluate the microkinetic behavior of a reaction network for which mainly two different approaches exist. These are the mean field approximation (MFA) and kinetic Monte Carlo (KMC), and depending on the problem at hand, each method has its advantages and disadvantages.[16] Another aspect is the complexity of the reaction network, which entails the considered elementary steps, interactions between adsorbates, as well as the different sites available on the catalyst surfaces. Neglecting an important factor, which is usually not known in advance, leads to wrong

results as the model does not represent the real mechanism and synergistic effects among adsorbates are missed. Yet, each step toward a more detailed computational description of a reaction network leads to a substantial amount of intricate computations necessary, which hinder the efficient screening of a large amount of different possible catalyst materials.[17] Especially, the energetics of adsorbates and transition states (TSs) are difficult to obtain. To this end, thermochemical relationships provide remedy, as it was found that adsorption energetics between a series of transition metals (TMs) scale linearly with respect to the kind of base element interacting with the surface. This can be exploited to restrict the necessary calculations with often reasonable accuracy.[18] Although such an approach works well for pure metal catalysts, it can fail for more complex catalyst materials.[2, 19]

In the search for improved catalysts, alloys are a promising group of candidates.[19] Optimal catalysts need to provide a balance between too strong and too weak adsorbate binding. Too strong adsorption leads to a poisoning of the catalyst surface, while too weak adsorption prevent surface reactions. Due to a variety of adsorption sites with different metal compositions, alloy surfaces hold the potential to provide superior adsorption strengths compared to pure metal facets. However, such mixed-metal materials often pose a problem for thermochemical relationships. Machine learning (ML) techniques might serve as a more flexible replacement for the reliable predictions of adsorption energetics. These data-driven approaches are similarly based on a set of known properties for the prediction of unknown ones, but the resulting description of the adsorption energy is more accurate and more general compared to the scaling relations.[19] The advantage of alloy catalyst materials is also its main drawback as the number of possibilities of different combinations is vast. Additionally, depending on the mixture ratio, different surface facets and crystal geometries are possible. However, there is a gleam of hope: There usually exist only a limited amount of stable or meta-stable bulk phases for a certain mixture ratio. To this end, databases which contain information about formation enthalpies of different mixture ratios and phases serve as an ideal starting point of further investigations.

Besides the efficient screening of possible novel catalyst materials and their active sites, a second important issue is the implications resulting from the chosen modeling approach. While the MFA only relies on a mean coverage, thereby neglecting the atomic scale resolution, KMC provides an in-depth view on the actual positions of the adsorbates.[16] Hence, KMC allows for the straightforward inclusion of lateral interactions among neighboring adsorbates. This effect is known to have a potentially decisive impact on the reaction energetics [20] and consequently the potential to render elementary steps accessible or inaccessible. Although there exist workarounds such as coverage-dependent reaction rates within the MFA,[21] these can not account for variations in the local coverages. On the other hand, modelling large reaction networks with long ranging interactions with standard KMC is computationally challenging. To this end, different schemes, i.e., temporal acceleration,[22, 23] ML,[19] and cluster expansions (CEs) [24] have been applied to overcome this limitation.

The above challenges called for the investigations published in Refs. 2 and 3 this cumulative thesis is centered around. In Ref. 2, data-driven material design methods for an efficient catalyst screening for the methanation reaction are applied. Relying on simple properties

of a bulk catalyst material, its surface, or the atomic species from literature and density functional theory (DFT), ML provides a means to determine a relation to the respective adsorption energy. The applied ML method Sure Independence Screening and Sparsifying Operator (SISSO) boils down to generating a simple linear descriptor of non-linear features using regularization schemes. Such schemes can be understood as a means for dimensionality reduction to only extract the most important features with the largest correlations to the quantity of interest – the adsorption energy. We used the so-obtained descriptor to predict the adsorption energies on binary alloys, which in turn served as inputs to mean-field microkinetic modeling (MKM) in order to evaluate the impact of the active site resolution on the prediction of possible improved methanation catalysts. We found that, especially for layered alloy structures, simpler microkinetic models significantly underestimate the turnover frequency (TOF) due to differences in the energetics of important intermediates. In Ref. 3, the focus is directed toward the importance of lateral interactions and the used MKM method to quantitatively reproduce experimental activity and selectivity trends for the CO hydrogenation toward higher oxygenate species. Comparing the predictions of corresponding MFA and KMC models, we found that the MFA – while providing reasonably accurate results under certain circumstances – will fail to predict the correct selectivity trends due to the lack of accurately described local changes in adsorbate coverage.

The following Chap. 2 is intended to provide a framework and overview to the theories and methods around the publications this thesis is based on. Sec. 2.1 starts with the CO hydrogenation and introduces the basics of establishing energetic data from DFT. In Sec. 2.2, the concept of compressed sensing (CS) together with the SISSO method is addressed. The following presentation of materials databases with their associated toolkits and in particular the Automatic Flow (AFLOW) framework sets the stage for a detailed discussion of the fundamental concepts of MKM in Sec. 2.3. Lastly, Chap. 3 presents brief summaries and the individual author contributions of the two articles, while Chap. 4 completes the thesis with a summary and outlook.

2 Theory

2.1 CO Hydrogenation: Experiment and Theoretical Modeling

The conversion of synthesis gas (syngas) – a mixture of carbon monoxide (CO) and hydrogen (H_2) – is a common platform for the production of value added products.[8, 25] Especially the synthesis of hydrocarbons and oxygenates for the use as fine chemicals or fuels from renewable feedstock is a great leap toward a more ecological value chain. Nowadays, syngas is mainly produced from fossil resources like gaseous or liquid carbohydrates (here saturated alkanes (C_nH_{2n+2})) and water (H_2O)

$$C_nH_{2n+2} + n\,H_2O \rightleftharpoons n\,CO + (2n+1)\,H_2 \quad , \tag{2.1}$$

or solid raw materials like coal, whose main constituent is C

$$C + H_2O \rightleftharpoons CO + H_2 \quad , \tag{2.2}$$

which renders its carbon footprint unfavorable.[26, 27] In order to make the transition toward sustainability it can also be produced from other feed stock such as biomass, plastic, and waste (cf. Eq. (2.1)) or carbon dioxide (CO_2)

$$CO_2 + H_2O \rightleftharpoons CO + H_2 + O_2 \quad . \tag{2.3}$$

Subsequent utilization of the formed syngas usually requires a specific CO/H_2 ratio.[8] This ratio can be adjusted with appropriate conditions and catalysts via the water-gas shift reaction

$$CO + H_2O \rightleftharpoons H_2 + CO_2 \quad . \tag{2.4}$$

Within the Fischer-Tropsch process, the methanation reaction, methanol production, or ethanol production, the CO of the syngas is transformed and hydrogenated to different hydrocarbons and oxygenates.[8, 25, 28–46] Besides the dynamic interplay of reaction conditions and the H_2 activation ability, the main factor impacting product selectivity is the character of CO adsorption. Generally speaking, if the catalyst allows for dissociative CO adsorption, longer carbon chains are formed and if H_2 is easily dissociated, hydrogenation reactions will occur more frequently. The CO activation ability is correlated with the location

of the catalyst transition metal (TM) in the periodic table of elements. Dissociative adsorption occurs in the left region leading to products such as methane (CH_4), higher alcohols, or C_{2+} oxygenates, while non-dissociative adsorption in the right, more noble region, e.g., Cu, mainly produces methanol.[8, 47, 48]

Tuning, enhancing, and optimizing catalytic activity and product selectivity is one of the major objectives in syngas conversion. One possibility is to make use of synergistic effects due to mixing of different TMs to form alloys. In this work, one focus lies on the identification of improved bimetallic methanation catalysts and the effect of bifunctional gains (*vide infra*) within this reaction network due to the introduction of different active sites, which have been reported to have a large impact on other reaction networks.[2, 8, 49, 50] The second focus lies on the detailed modeling of syngas conversion to CH_4, acetaldehyde, and ethanol on Rh catalysts. Shedding light on the microscopic scale and its connection and influence on the macroscopic scale gives invaluable mechanistic understanding and insights into the disputed cause of the observed structure activity relationships. The microscopic scale is represented by the active sites, their individual occupation with adsorbates, as well as the interaction between adsorbates. The macroscopic scale is represented by the observable activity and selectivity.[21, 51] In the following, the theoretical foundations toward the computational modeling of these two topics are introduced together with the related literature.

2.1.1 Density Functional Theory

The method of choice to obtain an understanding of the electronic structure of chemical systems – the atomistic scale – is density functional theory (DFT). Its advantage over Hartree–Fock and post-Hartree–Fock methods is that the computational cost is lower. However, this comes at a price: the exchange and correlation energy of the electrons are unknown and only approximations to their true value exist. According to Jacob's ladder of DFT functionals, they are arranged in rungs with an increasing level of accuracy: local density approximation (LDA), generalized gradient approximation (GGA), meta-generalized gradient approximation, hybrid functionals, and functionals including many-body terms.[52]

Besides the choice of the functional, the choice of basis functions to represent the electron density is an important factor and the best option depends on the problem.[53] The three main basis sets are a localized, atom centered basis, a real-space basis, or plane-waves.[53, 54] A localized basis set is usually a good choice for atoms or molecules, while periodic structures like crystals or surfaces are best represented in a plane-wave basis. To avoid an excessive amount of basis functions within a plane-wave approach due to the electrons close to the nucleus, these electrons are modeled using effective pseudopotentials, which decrease the computational burden. As the focus of this work lies on catalyst surfaces, we used the QUANTUM ESPRESSO plane-wave code [55] in combination with ultra soft pseudopotentials.

One of the problems of commonly used functionals is that they have troubles with modeling dispersion interactions.[56] To this end, several schemes exist to include dispersive van-der-Waals interactions and to correct for these shortcomings. One example of such a

functional is the Bayesian error estimation functional with van der Waals correlation (BEEF-vdW). Additionally, compared to the traditional approaches for developing functionals based on simple fitting procedures to experimental data or constraints on analytic properties, the development used ideas from Bayesian statistics and machine learning (ML) methods in the fitting to avoid overfitting and yield a general-purpose exchange-correlation approximation to reliably describe weak and strong interactions. This approach also provides an error estimate by determining the standard deviation for an ensemble of functionals around the fit parameters. Throughout this thesis, the BEEF-vdW functional was used for the determination of adsorption energies and in Ref. 3 the general shortcoming that GGA functionals tend to overbind CO on TM surfaces was investigated.[57–59]

2.1.2 Formation Energy Approach and Adsorption Energetics

The DFT-derived energy required or released upon an adsorption process is determined via the formation energy approach. The basis to this approach is a common gas phase reference energy set consisting of an energetic reference R_i of each atomic species i involved in the adsorption process to maintain thermodynamic consistency.[15] In order to calculate the adsorption process of CO to a surface, a reference state for C and O is required. The choice can be arbitrary but a common reference to this end are molecular H_2, CH_4, and H_2O. The atomic references R_i are then calculated as

$$R_\mathrm{H} = \frac{1}{2}E_{\mathrm{H_2}} \quad , \quad R_\mathrm{C} = E_{\mathrm{CH_4}} - 2E_{\mathrm{H_2}} \quad \text{, and} \quad R_\mathrm{O} = E_{\mathrm{H_2O}} - E_{\mathrm{H_2}} \quad , \tag{2.5-2.7}$$

where E_x is the energy of the molecule x as it is calculated by the DFT code. With these gas phase references, the formation energy E_x^form of an adsorbate x can be calculated using

$$E_x^\mathrm{form} = E_x^\mathrm{slab+ads} - E^\mathrm{slab} - \sum_{i \in x} n_{x,i} R_i \quad , \tag{2.8}$$

with the total energy of the adsorbed species on the slab $E_x^\mathrm{slab+ads}$, that of the clean slab E^slab, and the occurrence $n_{x,i}$ of atom i in the adsorbate x.

2.1.3 Free Energy and Vibrations

In a chemisorption process, the reactant molecule goes from the gas phase to adsorb to a solid by forming a bond to the catalyst active site. The energetics obtained from DFT only represent the electronic energy at $T = 0\,\mathrm{K}$ and $p = 0\,\mathrm{bar}$. For the effect of finite temperatures and pressures, the free energy of the gas phase molecules and adsorbates needs to be determined. The most simple approximation for gaseous species is the ideal gas approximation in which the individual particles do not interact and all degrees of freedom (DOFs) are independent of each other.[60, 61] The Gibbs free energy of such a particle is calculated via

$$G = H - TS = U + pV - TS \quad , \tag{2.9}$$

with the enthalpy H, temperature T, entropy S, internal energy U, pressure p, and volume V. The enthalpy H in turn is obtained via

$$H = E + E^{\mathrm{ZPE}} + \int_0^T C_p \, dT \quad , \tag{2.10}$$

with the electronic energy E, the zero-point energy E^{ZPE}, and the heat capacity at constant pressure C_p. The latter two terms require knowledge about the vibrational frequencies of the gaseous species. These can be obtained from DFT by evaluating the force constants as derivatives of the forces around the optimized structure. With the calculated vibrational frequencies, the zero-point vibrational energy is calculated by

$$E^{\mathrm{ZPE}} = \sum_i^{N^{\mathrm{modes}}} \frac{h\nu_i}{2} \quad , \tag{2.11}$$

with the Planck constant h and the vibrational frequency ν_i of mode i. For N atoms, the number of modes N^{modes} is $3N - 5$ for a linear species and $3N - 6$ for a nonlinear species. The heat capacity at constant pressure is

$$C_p = k_{\mathrm{B}} + C_V \quad , \tag{2.12}$$

with the Boltzmann constant k_{B} and the heat capacity at constant volume C_V. This, in turn, is comprised of a translational, a rotational, a vibrational, and an electronic part:

$$C_V = C_V^{\mathrm{trans}} + C_V^{\mathrm{rot}} + C_V^{\mathrm{vib}} + C_V^{\mathrm{el}} \quad . \tag{2.13}$$

Their contributions are

$$\int_0^T C_V^{\mathrm{trans}} \, dT = \frac{3}{2} k_{\mathrm{B}} T \quad , \tag{2.14}$$

$$\int_0^T C_V^{\mathrm{rot}} \, dT = \begin{cases} 0 & \text{(monatomic)} \\ k_{\mathrm{B}} T & \text{(linear)} \\ \frac{3}{2} k_{\mathrm{B}} T & \text{(nonlinear)} \end{cases} \text{, and} \tag{2.15}$$

$$\int_0^T C_V^{\mathrm{vib}} \, dT = \sum_i^{N^{\mathrm{modes}}} \frac{h\nu_i}{e^{\frac{h\nu_i}{k_{\mathrm{B}} T}} - 1} \quad , \tag{2.16}$$

while the electronic component is assumed to be zero. The entropy S of an ideal gas similarly contains a translational, rotational, vibrational, and electronic contribution at pressure p and a reference pressure p°:

$$S = S^{\mathrm{trans}} + S^{\mathrm{rot}} + S^{\mathrm{vib}} + S^{\mathrm{el}} - k_{\mathrm{B}} \ln\left(\frac{p}{p^\circ}\right) \quad . \tag{2.17}$$

Each contribution is calculated by the following expressions:

$$S^{\text{trans}} = k_{\text{B}} \left\{ \ln \left[\left(\frac{2\pi m k_{\text{B}} T}{h^2} \right)^{\frac{3}{2}} \frac{k_{\text{B}} T}{p^{\circ}} \right] + \frac{5}{2} \right\} \quad , \tag{2.18}$$

$$S^{\text{rot}} = \begin{cases} 0 & \text{(monatomic)} \\ k_{\text{B}} \left[\ln \left(\frac{8\pi^2 I k_{\text{B}} T}{\sigma h^2} \right) + 1 \right] & \text{(linear)} \\ k_{\text{B}} \left\{ \ln \left[\sqrt{\frac{\pi I_a I_b I_c}{\sigma}} \left(\frac{8\pi^2 k_{\text{B}} T}{h^2} \right)^{\frac{3}{2}} \right] + \frac{3}{2} \right\} & \text{(nonlinear)} \end{cases} , \tag{2.19}$$

$$S^{\text{vib}} = k_{\text{B}} \sum_{i}^{N^{\text{modes}}} \frac{\frac{h\nu_i}{k_{\text{B}} T}}{\exp\left(\frac{h\nu_i}{k_{\text{B}} T}\right) - 1} - \ln\left(1 - \exp\left(-\frac{h\nu_i}{k_{\text{B}} T}\right)\right) \quad \text{, and} \tag{2.20}$$

$$S^{\text{el}} = k_{\text{B}} \ln(2s + 1) \quad . \tag{2.21}$$

Here, m is the mass of the particle, I_a, I_b, and I_c the principle moments of inertia of a nonlinear molecule, I the moment of inertia of a linear molecule, σ the symmetry number, and s the total spin. With these contributions, the free energy in Eq. (2.9) for an ideal gas can be determined.[60, 61]

A similar approach can be employed for adsorbates by assuming that the translational and rotational DOFs get converted to vibrational DOFs upon chemisorption and treating them harmonically. Additionally, assuming the pV term of the enthalpy in Eq. (2.9) to be negligible leads to

$$G \approx U - TS = E + E^{\text{ZPE}} + \int_0^T C_V^{\text{vib}}\, dT - TS^{\text{vib}} \quad . \tag{2.22}$$

Although this approximation is quite rough, as the associated vibrations are hindered translations or rotations, it is widely and successfully applied as a good estimate.[60, 61]

2.1.4 Phase Stability

The thermodynamic stability of a certain bulk phase is determined by the Gibbs free energy G (Eq. (2.9)) of the material.[62] For a binary phase $A_{x_A}B_{x_B}$ to be stable, its formation Gibbs free energy ΔG needs to be exergonic:

$$\Delta G = G_{A_{x_A}B_{x_B}} - (x_A G_A + x_B G_B) < 0 \quad , \tag{2.23}$$

with the atomic fraction x_A (x_B) of element A (B) and the free energies $G_{A_{x_A}B_{x_B}}$ of the alloy as well as of its components G_A and G_B. At zero temperature, the Gibbs free energy equals the enthalpy:

$$H_{\text{f}} = H_{A_{x_A}B_{x_B}} - (x_A H_A + x_B H_B) \quad , \tag{2.24}$$

with the enthalpies $H_{A_{x_A}B_{x_B}}$ of the alloy and its components H_A and H_B. Neglecting vibrational contributions, which is a valid assumption for heavier elements such as TMs, the energy obtained from DFT is equivalent to the formation enthalpy H_f. The normalization of the alloy formation enthalpies by the enthalpies of the stable pure elemental phases gives a measure of relative stability. While a positive value indicates a thermodynamically unstable phase, negative values reflect stable and metastable phases.[62]

2.1.5 Transition States and How to Obtain Them

Finding transition states (TSs) is an essential but demanding task for determining energetic barriers for chemical reactions.[61] They do not represent a minimum on the potential energy surface (PES) but rather reflect a saddle point along the minimum energy path (MEP) from an initial state (IS) to a final state (FS), which is discerned by a maximum in one dimension of the PES and corresponds to a Hessian with only a single negative eigenvalue. Consequently, well-known minimization algorithms are not applicable and two widely employed examples of algorithms to locate saddle points are the dimer method [63] and the nudged elastic band (NEB) [64, 65] method, in particular in its climbing image (CI) variant.[66] The dimer method only requires an IS as starting point, from which it iteratively follows the lowest curvature mode until it reaches a saddle point. On the other hand, the NEB method needs both, IS and FS, defined. Interpolating between the two states and defining a number of intermediate structures (images) along this interpolation yields a starting point. The different images are then relaxed, tied together by a spring constant. The one image closest to the maximum will go upwards on the PES in the CI method, making sure the segment around the maximum is sufficiently represented and this image ends up close the maximum. Hence, the energies of TSs and activation barriers can be calculated, and the contributions to the free energy can be determined in the same way as for adsorbates, with the only difference that the one imaginary mode refers to the reaction coordinate.[61]

2.1.6 Linear Scaling Relationships and Transition State Scaling

DFT calculations for the required energetics of a complex reaction network or different catalyst materials are the limiting factor for an *ab initio* study.[67] To circumvent this issue, these energetics can be predicted with semi-empirical linear scaling relationships,[68] which exploit the observation that adsorbates and intermediates with the same binding atom or even different ones respond similarly to changes in the metal surface. These relationships are based on the *d*-band model [18, 69] as the chemisorption strength depends on the *d*-band center of the TM. For hydrogenation reactions such relationships have been extensively studied and successfully applied for the determination of volcano plots, which lead to a drastic reduction of first-principles (1p) calculations. These plots represent the adsorption energy (or energies) of an adsorbate at an active site of different materials as descriptor that leads to the optimal catalyst activity. For the case of AH_x (A = C, N, O) such a relationship is of the form

$$\Delta E^{\text{ads}}_{\text{AH}_x} = \gamma \Delta E^{\text{ads}}_{\text{A}} + \xi \quad , \tag{2.25}$$

with the adsorption energy of the atomic adsorbate $\Delta E_{\mathrm{A}}^{\mathrm{ads}}$, and the fit constants γ and ξ.[11, 67, 68, 70, 71]

Similarly, such relationships exist for TSs. The so-called Brønsted-Evans-Polanyi (BEP) relationships correlate the activation energy E_{a}, i.e., the difference between the energy of the TS and that of the IS, with the reaction energy ΔE_{rxn}, i.e., the difference between the energy of the FS and that of the IS:[67, 72–74]

$$E_{\mathrm{a}} = E_{\mathrm{TS}} - E_{\mathrm{IS}} = n_{\mathrm{BEP}} \Delta E_{\mathrm{rxn}} + m_{\mathrm{BEP}} = n_{\mathrm{BEP}} (E_{\mathrm{FS}} - E_{\mathrm{IS}}) + m_{\mathrm{BEP}} \quad , \tag{2.26}$$

with the fit constants n_{BEP} and m_{BEP}. Additionally, TS scaling relations correlate the TS energy with the IS or FS.[67, 75] Depending on the similarity of the TS to the reactants or the products, it was found that different linear scaling formulations perform differently accurate. For a reactant-like TS the best linear free energy relation will be obtained by correlating the TS energy with the IS energy together with the constants n_{IS} and m_{IS}:[76]

$$E_{\mathrm{TS}} = n_{\mathrm{IS}} \cdot E_{\mathrm{IS}} + m_{\mathrm{IS}} \quad . \tag{2.27}$$

If the TS is product-like, scaling with the FS energy yields the best correlation:

$$E_{\mathrm{TS}} = n_{\mathrm{FS}} \cdot E_{\mathrm{FS}} + m_{\mathrm{FS}} \quad , \tag{2.28}$$

with the fit constants n_{FS} and m_{FS}.

2.1.7 Lateral Interactions

Interactions between adsorbed species are often neglected in microkinetic studies due to large computational requirements attached to their inclusion.[19, 77] However, such interactions can have a decisive impact on the modeling result.[24, 78–81] These differences arise due to changes in the underlying PES and, consequently, the energetic barriers depending on the individual neighboring adsorbates. Such interactions may stabilize or destabilize an adsorbed species and alter the energetics of a process in various nuances, possibly leading to island formation (attractive) or uniform adlayers (repulsive) due to maximization of adsorbate distances. Since the energetic term enters the transition probability in the exponent (*vide infra*), even small variations can have a large impact on the whole reaction network.[19, 82]

Computationally, such energetic effects within a mean field approximation (MFA) approach can be taken into account with a coverage dependent adsorption energy function. To this end the formation energies of different coverages are fit to an, e.g., piece-wise linear function.[21] However, because of the approximation of an average coverage, this can lead to inaccuracies due to the neglect of local differences in surface configurations.[3, 24, 83, 84]

Within kinetic Monte Carlo (KMC), lateral interactions can be modeled using a cluster expansion (CE) approach of a lattice-gas Hamiltonian.[85–88] For a single adsorbate and one

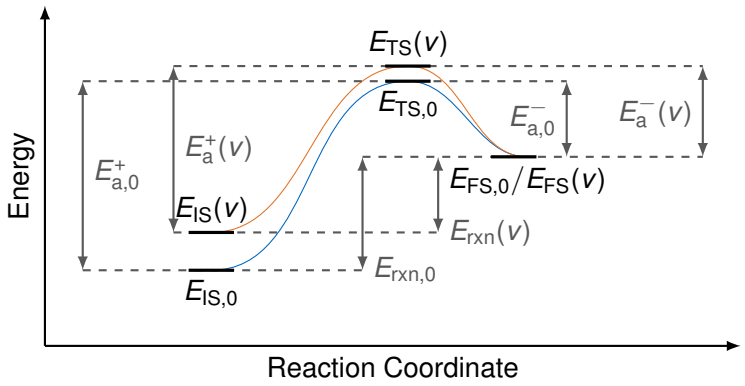


Figure 2.1: Energy profiles of a reaction coordinate at the zero coverage limit (blue) and with repulsive lateral interactions (orange) with the corresponding energetics for the calculation of the quantities in Eqs. (2.30) to (2.33). Adapted from *J. Chem. Phys.* 2013, 139, 224706 under the terms of the CC BY 3.0 license. © 2013 Nielsen, d'Avezac, Hetherington, and Stamatakis. Published by AIP Publishing.

site, the individual energetic contributions due to lateral interactions are fit to a set of DFT calculations with varying amounts of adsorbed species using the ansatz

$$E(v) = E\sum_{s} n_s + \sum_{a=1}^{b} V_a^p \sum_{(s<t)_a} n_s n_t + \sum_{a=1}^{c} V_a^t \sum_{(s<t<u)_a} n_s n_t n_u + \ldots \quad , \tag{2.29}$$

where $E(v)$ denotes the energy of lattice state v, which is determined by the sum of the binding energy E, the pairwise interaction energies V_a^p, and the triplet interaction energies V_a^t of the adsorbates on the sites s, t, and u. The n represent the occupation number on the sites with one of the adsorbates and can have the values 0 or 1. Considering a total of b different pair interaction patterns and c different triplet interaction patterns on the ath nearest neighbor sites, the summation over the sites adds each unique pattern only once if the occupied sites form such a pattern. Larger interaction ranges are incorporated in the same way,[77, 88] but typically CEs only consider nearest neighbor interactions, especially for more complex models, in order to keep the computational requirements manageable.[89, 90] TSs pose an additional challenge when including lateral interactions. In order to determine the interaction parameters (Eq. (2.29)) a set of TS calculations with different adsorbates and different coverages need to be performed. However, as this is mostly out of reach except for simple reaction mechanisms the effect of lateral interactions on TS energies are approximated using BEP relations according to Nielsen et al.[91] (cf. Fig. 2.1):

$$E_{rxn}(v) = E_{FS}(v) - E_{IS}(v) \tag{2.30}$$

and

$$E_{rxn,0} = E_{FS,0} - E_{IS,0} \quad , \tag{2.31}$$

with the reaction energy $E_{\mathrm{rxn}}(\nu)$ ($E_{\mathrm{rxn},0}$), the energy of the IS $E_{\mathrm{IS}}(\nu)$ ($E_{\mathrm{IS},0}$), and the energy of the FS $E_{\mathrm{FS}}(\nu)$ ($E_{\mathrm{FS},0}$) at state ν (zero coverage). The energies of interest are the activation energies $E_{\mathrm{a}}^{+}(\nu)$ and $E_{\mathrm{a}}^{-}(\nu)$ of the forward and backward reaction, respectively:

$$E_{\mathrm{a}}^{+}(\nu) = \max(0, E_{\mathrm{rxn}}(\nu), E_{\mathrm{a},0}^{+} + n_{\mathrm{BEP}}(E_{\mathrm{rxn}}(\nu) - E_{\mathrm{rxn},0})) \tag{2.32}$$

and

$$E_{\mathrm{a}}^{-}(\nu) = \max(0, -E_{\mathrm{rxn}}(\nu), E_{\mathrm{a},0}^{-} + (1 - n_{\mathrm{BEP}})(E_{\mathrm{rxn}}(\nu) - E_{\mathrm{rxn},0})) \quad , \tag{2.33}$$

with the BEP factor n_{BEP}, which is bound by $[0, 1]$ and is a measure of how reactant-like (0) or product-like (1) the TS is. This factor is determined by fitting the activation energy of different lattice states ν to $E_{\mathrm{rxn}}(\nu) - E_{\mathrm{rxn},0}$.

2.2 Machine Learning

With the constant increase of computational resources and the concomitant ability to generate larger and larger amounts of data, the field of cheminformatics – among all other sciences – started to enter the 4$^{\mathrm{th}}$ paradigm.[92–94] Entailing the first three paradigms of experiment, theory, and simulation, it describes data-driven science. In combination with artificial intelligence (AI) and its sub-field machine learning (ML), it was then also possible to handle the vast amount of big-data, which represents the key for all ML techniques to extract and predict valuable information and patterns from the spaces spanned by the data *in silico*. The amount of data, on the other hand, imposes tight requirements on, e.g., the required storage space, the heterogeneity, or the organization of the database. Efforts to combine and create searchable databases for different applications and properties are, therefore, undertaken and still ongoing in various fields of science, public administration, and industry.[95–99]

ML is mainly classified in four groups on how an algorithm obtains the final model or descriptor.[100–102] The first type is called supervised learning in which the algorithm learns from known examples. To this end, it is provided with a data set containing the desired inputs and outputs, commonly referred to as labeled data. The outputs represent certain properties and the inputs are features to describe these properties by finding an accurate relationship to predict unknown outputs or patterns. Common applications of supervised learning algorithms, i.e., linear regression, support vector machines, or neural networks, are classification for discrete outputs or regression for continuous outputs.[100] If only features serve as input and commonalities are subject to identification it is called unsupervised learning: the algorithm needs to determine patterns within the presented unlabeled data for the purpose of, e.g., clustering, dimensionality reduction, or association. Common algorithms for this task are principal component analysis or K-means clustering.[102] The so-called semi-supervised learning falls in between the first two types. Both, labeled and unlabeled data is used and the algorithm is trained to classify the unlabeled data and provide a better model than solely relying on supervised learning. In the fourth category – the reinforcement

learning – the algorithm performs a recurrent dynamic feedback loop to maximize a reward, which involves an action on the presented data and monitoring the response.[100–102]

This work is based on the supervised learning algorithm Sure Independence Screening and Sparsifying Operator (SISSO) from the field of compressed sensing (CS) introduced in Sec. 2.2.1 for the prediction of adsorption energies.[103, 104] For a supervised approach, the database, which is used to train the ML algorithm and derive a mathematical descriptor, needs to be representative for the entire chemical space of interest to make accurate predictions. Not performed properly, this might lead to overfitting, which means that the model will be accurate within the provided set but not for unseen data. Therefore, a valid approach is to set aside a test set from all available samples and repeatedly split the remaining data into a training and a validation set. The model is built upon the training set, evaluated and refined on the validation set and finally evaluated on the test set by means of comparing a suitable metric like the root mean squared error (RMSE) or mean absolute error (MAE) for a continuous quantity or, e.g., precision for classification models. In a next stage, such a forward approach to predict certain properties can be reversed to obtain models for the discovery of materials from a set of optimal properties.[94]

2.2.1 Compressed Sensing

CS originates from the field of signal processing. Given a small set of measurements, the aim of CS is to recover the original signal by identifying its characteristic components.[2, 103] Mathematically speaking, this represents an under-determined system of linear equations

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \quad , \tag{2.34}$$

with the coefficient or sensing matrix $\boldsymbol{A}$, the measurements $\boldsymbol{b}$, and the unknowns $\boldsymbol{x}$. Each row in $\boldsymbol{A}$ holds the different hypothetical components or features of a signal sample. As there are infinite solutions to such a system, CS requires sparse solutions, i.e., a small number of non-zero entries in $\boldsymbol{x}$, representing the characteristic components. Effectively, this yields a dimensionality reduction. Up until now, everything is linear algebra. Non-linearity can be included by creating candidate features, which are non-linear functions.[2, 103]

Sure Independence Screening and Sparsifying Operator

SISSO constructs the feature space $\boldsymbol{\Phi}_n$ of rung n, which are the hypothetical components contained in $\boldsymbol{A}$, based on a hierarchical approach by iteratively combining the features ϕ of the previous iteration (rung) n starting from assumed primary features via

$$\boldsymbol{\Phi}_n \equiv \bigcup_{i=1}^{n} \hat{\boldsymbol{H}}^{(m)} [\phi_1, \phi_2] , \quad \forall \phi_1, \phi_2 \in \boldsymbol{\Phi}_{n-1} \tag{2.35}$$

to generate a wide range of possible components and finally arrive at models with better accuracy. The size of the feature space grows rapidly depending on the size and choice of the (non-linear) mathematical operators in the operator set

$$\hat{\boldsymbol{H}}^{(m)} \equiv \{I, +, -, \times, \div, \exp, \log, \sqrt{}, ...\} \tag{2.36}$$

Class	Name	Abbreviation	Unit
Atomic	Electron affinity	EA	eV
	Ionization potential	IP	eV
	Pauling electronegativity	PE	–
Bulk	Coupling matrix element squared	V_{ad}^2	–
	Nearest neighbor distance	$bulk_{nnd}$	Å
	Radius of *d*-orbitals	r_d	Å
Site	Coordination number	CN	#atoms
	d-band center	ε_d	eV
	d-band filling	f_d	#states
	d-band width	W_d	eV
	Density of *d*-states at Fermi level	DOS_d	eV^{-1}
	Density of *sp*-states at Fermi level	DOS_{sp}	eV^{-1}
	Nearest neighbor distance	$site_{nnd}$	Å
	Number of atoms in ensemble	$site_{no}$	#atoms
	sp-band filling	f_{sp}	#states
Surface	Work function	W	eV

Table 2.1: Properties of a clean catalyst surface used as primary features for the feature construction within SISSO as in Ref. 2.

as well as the user-defined rung n. As features can be quantities carrying units, only appropriate feature combinations are subject to binary operators as indicated by the superscript m.[103]

To obtain the best sparse solution to the linear system in Eq. (2.34), SISSO adds the regularization term $\lambda\|\boldsymbol{x}\|_0$ and minimizes with respect to $\boldsymbol{x}$ via

$$\arg\min_{\boldsymbol{x}} \left(\|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_0\right) \quad . \tag{2.37}$$

$\|\boldsymbol{x}\|_0$ is the ℓ_0 "norm", which is the number of non-zero entries in $\boldsymbol{x}$ and the parameter λ controls this number of entries within $\boldsymbol{x}$. Within the SISSO framework this regularization is the so-called sparsifying operator. As the ℓ_0 "norm" regularization is non-deterministic polynomial-time hard,[1] the feature space is previously reduced to a user-defined value by applying sure independence screening, i.e., pre-screening the features with the largest correlation to the quantity of interest. The larger this value, the higher the probability of finding the best descriptor, however, at the expense of an increase in computational cost. A logical extension to this algorithm termed multi-task learning is to jointly learn from multiple different input sets $\boldsymbol{A}$ simultaneously and then arriving at a joint solution containing all mutually important features. Hence, the generalization of the identified model is increased and the risk of overfitting is decreased compared to the single-task.[2, 103, 104]

[1] The computational cost grows exponentially with the size of the feature (sub)space rendering regularization impossible for large such spaces.

The study on the identification of improved methanation catalysts published in Ref. 2 investigates the influence of the active site representation within microkinetic modeling (MKM) (cf. Sec. 2.3). The missing link to obtain a descriptor for the prediction of adsorption energies is a training set consisting of primary features and the corresponding adsorption energetics as well as the retrieval of possible candidate catalyst surfaces. For the training set, we used and extended the features and density functional theory (DFT)-calculated database from Ref. 19 (cf. Tab. 2.1). We then fit the respective adsorption energetics using multi-task SISSO with each of the different adsorbates assigned to a single task. The primary features list only contains properties of the clean catalyst surfaces, in order to obtain a general and adsorbate-independent descriptor for the prediction of the energetics on alloy candidate catalysts. How these candidate catalysts were obtained is the topic of the following.

2.2.2 The AFLOW Framework

Large repositories consisting of a variety of *ab initio* data from DFT and high-throughput methods provide the ideal starting point for screening studies. Examples of such databases are the Materials Project,[105, 106] the Novel Materials Discovery (NOMAD) repository,[98] the Open Quantum Materials Database (OQMD),[107, 108] the Citrination database,[109] the Computational Materials Repository (CMR),[110] or the Automatic Flow (AFLOW) database.[111] Often, these databases are combined with automated generation, analysis, and characterization tools to provide an independent framework, which allows for the implementation of new structure discovery and optimization applications. In our search for improved methanation catalysts, we aim to identify suitable crystal phases of binary transition metal (TM) alloys to predict their respective adsorption energetics on the various sites on their exposed facets.[2] The reason why we focus on such binary alloys is that they hold the potential to break the scaling relations, which linearly correlate binding energies with corresponding activation energies, by introducing a bifunctionality through the mixed-metal site types. These relations were initially identified for various TMs as they are usually well performing catalysts for a variety of different reactions. However, only a few TMs fall into the region of an ideal catalyst in which the binding energy is large enough to bind the adsorbate while not hindering product formation.[49, 50, 112] The different adsorption energetics can enhance catalytic activity via synergistic effects. In a volcano plot, such materials pose "outliers" and are interesting candidates for further experimental investigations.

The AFLOW Database

As of July 2022 this database contains – besides properties of ternary and quaternary systems – properties of about 1,700 binary bulk crystal systems.[113] The information on the respective space group defining all possible symmetry operations of the lattice forms the basis of our screening study.[2] Hence, the AFLOW framework is especially suited for this purpose as it contains data on the formation enthalpies of such crystal phases and an application programming interface for their straightforward retrieval.[62, 111, 114, 115]

From a computational standpoint, alloys pose a big problem: which mixing ratios are thermodynamically stable and what is the arrangement of the elements within the bulk, that

translates to the space group the crystal structure belongs to. There are a total of 230 different space groups and combining two TMs enlarges the range of possible space groups dramatically compared to pure TMs. From these 230 unique groups 27 (with the numbers 168-194) are of the hexagonal and 36 (with the numbers 195-230) are of the cubic crystal system. Additionally, also meta-stable, i.e., kinetically "trapped" crystal structures might play a role. In a pure solid the face-centered cubic (fcc) structure corresponds to the space group 225 and is representative for most of the TMs within the training set. One exemption is Co, which crystallizes in a hexagonal close packed (hcp) structure with the space group 194, but the energetic difference to an fcc lattice is quite small so that at standard conditions both lattices may coexist. Another exemption is Fe, which crystallizes in a body-centered cubic (bcc) lattice, which has the space group 229.[116]

As a first step toward binary alloy catalyst screening, we performed a database search of a variety of different TM combinations in order to find fcc and bcc like (meta)stable bulk phases. The plethora of obtained information on different TM combinations and their respective space groups within the repository required a thorough pre-screening of the *ab initio* phase stability data.

Convex Hull

With the convex hull (CHULL) module of AFLOW,[62] phase diagrams of the bulk formation enthalpies H_f of binary alloys contained in the database can be generated. Figure 2.2 shows such a diagram of the binary Fe-Pt system. Each of the crosses and dots represent the formation enthalpy of an alloy within the repository on the ordinate with a certain mixing ratio on the abscissa. Referenced to the pure metals as the zero points, all phases with a negative H_f are either stable or metastable. The phases marked with dots are the stable ones at thermodynamic equilibrium, which define the convex hull (lines). The crosses mark metastable phases and depending on the kinetics of their formation they might reconstruct and/or separate into the closest thermodynamically stable phases.

The next step is the generation of the surface facets. As the SISSO training set contains the fcc(211) facet, we require space groups that show the same or a similar facet. Depending on the mixing ratio and the symmetry of the alloy, each direction as well as multiples of the *hkl* cut might expose a different atomic arrangement. From the different phases obtained from the convex hull data, we chose the two fcc space groups 221 and 225, and the two hcp space groups 187 and 164, which led to two or four different slab models per alloy phase. The respective adsorption energetics from the SISSO descriptor were subsequently used in combination with Brønsted-Evans-Polanyi (BEP) relations [72, 73] and transition state (TS) scaling [75] within MKM (cf. Sec. 2.3) to predict and compare the catalyst activity from models with increasing active site resolution.[2]

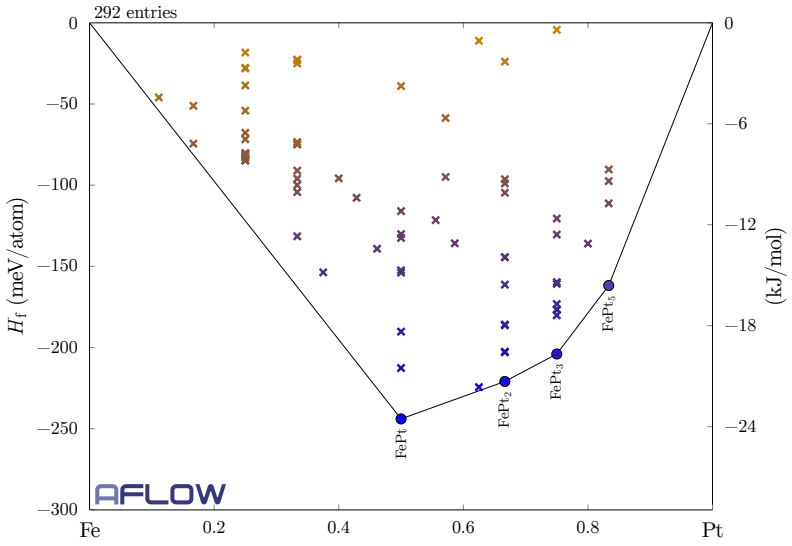


Figure 2.2: Convex hull plot of the binary Fe-Pt system as in Ref. 2 and as obtained by AFLOW-CHULL.[62] Phases of different mixing ratios on the convex hull line (dots) are stable, whereas the phases with formation enthalpies H_f between zero and the convex hull are metastable (crosses).

2.3 Microkinetic Modeling

In order to provide the framework for an analysis and to make predictions on the activity and selectivity of a catalyst within a computational approach, a microkinetic description is necessary. The overall process is modeled by a set of differential equations representing the elementary steps of a reaction mechanism – the master equation. In essence, this equation is a complete, high-dimensional description of a reaction network and no assumptions on important steps or the coverage and spatial arrangement of the adsorbed species are made. To perform a simulation the thermodynamic and kinetic parameters are required, i.e., the energetics of the adsorbates and transition states (TSs) involved in the mechanism to be modeled. In bottom-up multiscale modeling of the kinetics of catalytic processes, the most common tool for the determination of these parameters is density functional theory (DFT).[117] Explicitly solving the components of the master equation is at least a computationally very demanding task – especially if interactions between the different adsorbed species and surface nonidealities are included. These correlations are often neglected and idealized surface facets are used to avoid computational overhead. The master equation is then solved using the mean field approximation (MFA) (cf. Sec. 2.3.5) or kinetic Monte Carlo (KMC) (cf. Sec. 2.3.6). The so-obtained results give valuable insights into the reaction mechanism, structure-property relationships, and assist in reactor design. However, a word of caution must be mentioned here: depending on the true mechanism and the chosen microkinetic modeling (MKM) approach, the result may qualitatively deviate from the true so-

lution.[118] In the following, the above two main theories, by which heterogeneous catalytic reactions are modeled, and their prerequisites are introduced.[15, 19]

2.3.1 Reaction Networks

To describe a heterogeneously catalyzed chemical reaction on the microscopic scale, the overall reaction is subdivided into elementary steps forming a reaction network. This network encodes the elementary transformations of the reactants toward the products via all intermediates. A common example is the coupling of two species A and B into a third species C:

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C} \quad , \tag{2.38}$$

which can be represented by a Langmuir-Hinshelwood mechanism:

$$\mathrm{A} + * \underset{k_\mathrm{A}^-}{\overset{k_\mathrm{A}^+}{\rightleftharpoons}} \mathrm{A}^* \tag{2.39a}$$

$$\mathrm{B} + * \underset{k_\mathrm{B}^-}{\overset{k_\mathrm{B}^+}{\rightleftharpoons}} \mathrm{B}^* \tag{2.39b}$$

$$\mathrm{A}^* + \mathrm{B}^* \xrightarrow{k_\mathrm{C}} \mathrm{C} + 2* \tag{2.39c}$$

This sequence describes the adsorption of the two species A and B to the surface (Eqs. (2.39a) and (2.39b)) leading to the adsorbed species A^* and B^* with the rate constants of the forward direction k_A^+ and k_B^+ as well as the backward direction k_A^- and k_B^-, and a subsequent associative desorption of the product C (Eq. (2.39c)) with the rate constant k_C. For each of these elementary steps, the thermodynamic and kinetic parameters involved need to be determined. In combination with the reaction network, these inputs form the most basic information required, in order to perform microkinetic simulations.[117]

In an *ab initio* microkinetics study, the determination of the kinetic parameters is the bottleneck as it requires the most computational time. A common approach is to investigate preconceived mechanisms based on experimental results, heuristics, reaction rules, or readily available first-principles (1p) data. However, it is usually not known in advance whether a proposed step is important or not, i.e., whether this step plays a role in the true reaction mechanism. To make matters worse, one overall pathway might be the correct one for one catalyst material and certain reaction conditions but not for others, which poses a problem for screening studies. Omission of a path or reaction step might result in wrong activities and selectivities, and thus predict unsuitable catalysts.[117]

In order to avoid this eventuality and to make meaningful and precise predictions, it is important to consider all possible paths within the reaction network. With such an approach, no possible elementary reaction steps are missed. Only then, one can reliably decode the mechanism, focus on the dominant pathway,[2] and screen a large amount of possible catalyst materials. Such exhaustive reaction networks can be developed based on generation

[2] Assuming this pathway holds true for the other considered materials and conditions.

schemes relying on available data or reaction rules.[117, 119–123] Alternatively, the main reaction pathway can be generated adaptively on-the-fly without the need to define a fixed mechanism in advance. Based on the possible identified elementary steps from each lattice configuration, those with the lowest barriers are chosen. The reason for this is the high probabilities of occurrence associated with these processes. The drawback of such an approach, on the other hand, is its large computational cost, due to the vast number of 1p calculations needed. This can be reduced significantly by using approximate potentials for a first filtering of the possible transformations.[124–126] In this context, it is worth mentioning that only the rate-determining step (RDS) or the steps that are partly rate determining need to be calculated accurately (cf. Sec. 2.3.7) as they dictate the activity and selectivity. All other steps only play a subordinate role and inaccuracies do not have an effect on the overall outcome of the simulation. Since there is usually no prior knowledge, a hierarchical approach can significantly increase the simulation efficiency. Such an approach would start with a low level of accuracy and then iteratively improve the kinds of methods used only for the important transformations. Additionally, coarse graining by lumping less important steps together is an effective means to reduce the computational effort.[15, 117]

2.3.2 Lattice and Active Sites

Within a microscopic reaction network the intermediates bind to the surface at certain structural motifs, which correspond to local minima on the potential energy surface (PES) – the active sites. In Eqs. (2.39a) to (2.39c), this information is simply hidden behind the asterisk "$*$". Most studies assume a static crystalline surface with the active sites mapped onto a periodic lattice.[127–137] The connections between the different PES basins are the elementary steps. On an idealized surface, i.e., a face-centered cubic (fcc) (111) facet, there is only a small number of such motifs per unit cell, which are one on-top site, one bridge site, as well as the fcc and the hexagonal close packed (hcp) three-fold coordinated sites. However, species adsorbed to nearby sites influence the appearence of the PES. Due to this effect, the same elementary step is formally many different steps from an energetic point of view. Depending on the interactions, the energies of the initial state (IS), TS, and final state (FS) differ and lead to different probabilities of occurrence for the same transformation. Analogous to the discussion on reaction networks in Sec. 2.3.1, the active sites need to be identified and set in advance in order to determine the microscopic thermodynamic and kinetic parameters. One of the major challenges is that, under real operating conditions, there exists a large variety of such sites due to the structural complexity, i.e., the composition and geometry of the sites are dependent on the temperature and adsorbate concentrations. The latter are in turn dynamically dependent on the proceeding surface reactions and transport phenomena.[48]

In general, the larger the number of different active sites is, the higher is the number of possible reaction pathways and thus the simulation cost, primarily resulting from the concomitant number of 1p calculations. This poses a large problem for the simulation of real catalysts due to the mere number of possible active sites and the interplay between them. The lack of knowledge about the appearance as well as the effect of interactions between adsorbed

species themselves and their effect on the active sites are an ongoing field of research.[11, 15, 19, 117]

2.3.3 The Master Equation

The basis of MKM is to model chemical reactions as rare events. This entails the coarse-graining of vibrationally excited system states around the PES basins into metastable states, mapped to the aforementioned lattice sites. In this description, the dynamics is given by stochastic transitions between these discrete states. Those transitions are in particular adsorption, desorption, diffusion, and reactive processes. The coarse-graining of the intra-basin vibrational dynamics is, however, an approximation commonly employed and might not always hold.[138] Within a simulation, they do not contribute to the advancement of chemical transitions and – due to the considerably smaller time scale on which the movement takes place – would only introduce a major computational overhead. The long time span the system stays within such a basin vibrating around its minimum vindicates the assumption of a Markov process, i.e., the system forgot the preceding history rendering the inter-basin jumps independent of each other. The mathematical representation to find the system in any state v at time t is the master equation

$$\frac{dP_v(t)}{dt} = \sum_{w \neq v} k_{wv} P_w(t) - \sum_{w \neq v} k_{vw} P_v(t) \quad , \tag{2.40}$$

with the probability $P_v(t)$ of the system being in state v at time t and the transition probability rate k_{vw} of transitioning from state v to state w. The sum terms in the first sum express the transition rates of entering state v from any basin w, while the ones in the second sum express the respective rates of leaving state v toward any other basin w. For usual systems, however, the spanned space of the master equation is very high dimensional and this so-called *curse of dimensionality* precludes a straightforward solution.[19, 82, 118, 139]

2.3.4 Rate Constants

The transition probabilities or rate constants k_{vw} of all elementary steps need to be estimated to be used as input for MKM. The most common approach is to use typical empirical values fit to experimental results, which, however, might prevent an atomistic interpretation. For 1p modelling these probabilities are often based on DFT in combination with transition state theory (TST).[82, 140]

Within harmonic TST the rate constants for an exemplary process shown in Fig. 2.3 are approximated assuming harmonic vibrational modes leading to simple rate constant expressions of the form

$$k^{\mathrm{TST}}_{\mathrm{IS}\rightarrow\mathrm{FS}} = f^{\mathrm{TST}}_{\mathrm{IS}\rightarrow\mathrm{FS}} \frac{k_B T}{h} \exp\left(-\frac{E_a^+}{k_B T}\right) \quad , \tag{2.41}$$

with the activation energy in forward direction E_a^+ and the factor $f^{\mathrm{TST}}_{\mathrm{IS}\rightarrow\mathrm{FS}}$, which contains the

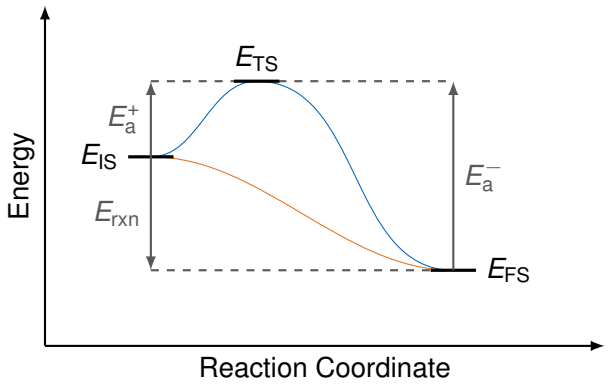


Figure 2.3: Energy profiles of a reaction coordinate for an activated (blue) and a non-activated reaction (orange) with the corresponding energetics for the calculation of rate constants.

partition functions Q of the IS and the TS:

$$f^{\mathrm{TST}}_{\mathrm{IS}\to\mathrm{FS}} = \frac{Q_{\mathrm{TS}}}{Q_{\mathrm{IS}}} \quad . \tag{2.42}$$

The factor $f^{\mathrm{TST}}_{\mathrm{IS}\to\mathrm{FS}}$ is a measure of the fraction of particles being reflected[3] due to deviations from the minimum energy path (MEP), which is the connection of two basins with the smallest energetic barrier. For more than one component in either state, these Q are the products of the individual partition functions. At steady-state, the probability for each state remains constant over time and the sum of all transitions into one state v is equal to the sum of transitions out of state v. Hypothetically turning off all processes except one single pair between two states w and v will lead to a relaxation into thermodynamic equilibrium between the two states, and the forward ($k_{wv}P_w$) and backward transition flows ($k_{vw}P_v$) balance each other. This is the so-called detailed balance and applies to each pair of forward and reverse transition within the reaction network. The probabilities P_w and P_v are proportional to the respective Boltzmann weights, which leads to the expression for the reversed process $k^{\mathrm{TST}}_{\mathrm{FS}\to\mathrm{IS}}$ via

$$\frac{k^{\mathrm{TST}}_{\mathrm{IS}\to\mathrm{FS}}}{k^{\mathrm{TST}}_{\mathrm{FS}\to\mathrm{IS}}} = \exp\left(\frac{G_{\mathrm{FS}} - G_{\mathrm{IS}}}{k_B T}\right) \quad , \tag{2.43}$$

with the free energy G_{IS} (G_{FS}) of the IS (FS). The small pV term from the enthalpy H in Eq. (2.9) is commonly neglected for adsorbed species. Furthermore, for reactions involving gaseous species, their free energy is replaced with the respective chemical potentials μ.[11, 15, 89, 132, 137, 141]

At the beginning of a catalytic cycle, an atom or molecule needs to adsorb to the catalyst surface. For an impinging gaseous species i onto the surface site s leading to adsorbate j

[3] The crossing of the barrier is considered a purely classical event. Quantum mechanical tunneling is not possible.

in the case of an activated process (cf. blue line in Fig. 2.3) the partition functions are

$$Q_{\mathrm{IS}} = q_i^{\mathrm{vib}}\, q_i^{\mathrm{rot}}\, q_i^{\mathrm{trans3D}} \tag{2.44}$$

and

$$Q_{\mathrm{TS}} = q_{\mathrm{TS}}^{\mathrm{vib}} \quad , \tag{2.45}$$

with the vibrational, rotational, and translational components. The TS is assumed to be close to the adsorbed state and, therefore, only contains vibrational components. For the approaching gas particle, its distance from the surface serves as the reaction coordinate. This particular degree of freedom (DOF) is thus separated within the three-dimensional translational partition function leading to

$$q_i^{\mathrm{trans3D}} = q_i^{\mathrm{trans2D}}\, d\, \frac{\sqrt{2\pi m_i k_B T}}{h} \quad , \tag{2.46}$$

with the mass m_i of species i and its distance d from the surface. Inserting Eqs. (2.42) and (2.44) to (2.46) into Eq. (2.41), and multiplying with the number N_i of molecules i in the gas phase, yields the rate constant for adsorption onto site s:

$$k_{i,s}^{\mathrm{ad}} = \frac{q_{\mathrm{TS}}^{\mathrm{vib}}}{q_i^{\mathrm{vib}}\, q_i^{\mathrm{rot}}\, q_i^{\mathrm{trans2D}}} \frac{N_i k_B T}{d\sqrt{2\pi m_i k_B T}} \exp\left(-\frac{E_a^+}{k_B T}\right) \quad . \tag{2.47}$$

Assuming an ideal gas ($p_i V = N_i k_B T$ and $V = dA$) impinging onto the active site leads to

$$k_{i,s}^{\mathrm{ad}} = \frac{q_{\mathrm{TS}}^{\mathrm{vib}}}{q_i^{\mathrm{vib}}\, q_i^{\mathrm{rot}}\, q_i^{\mathrm{trans2D}}} \frac{p_i A_s}{\sqrt{2\pi m_i k_B T}} \exp\left(-\frac{E_a^+}{k_B T}\right) = \tilde{S}_{i,s} \frac{p_i A_{\mathrm{uc}}}{\sqrt{2\pi m_i k_B T}} \quad , \tag{2.48}$$

with the the partial pressure p_i of species i, the surface area A_{uc} (A_s) of the unit cell (site s), temperature T, and the sticking coefficient

$$\tilde{S}_{i,s} = \frac{q_{\mathrm{TS}}^{\mathrm{vib}}}{q_i^{\mathrm{vib}}\, q_i^{\mathrm{rot}}\, q_i^{\mathrm{trans2D}}} \frac{A_s}{A_{\mathrm{uc}}} \exp\left(-\frac{E_a^+}{k_B T}\right) \quad . \tag{2.49}$$

This coefficient contains the statistical average of the positions, velocities, and internal DOFs of the impinging gaseous species i onto the surface and describes the fraction of how many particles successfully adsorb on a vacant site s. The free energy of a TS or adsorbate is given by

$$G_{i,s} = E_{i,s} - k_B T \ln(q_{i,s}^{\mathrm{vib}}) \tag{2.50}$$

and the chemical potential of a gaseous species by

$$\mu_i = E_{\mathrm{gas},i} - k_B T \ln\left(q_i^{\mathrm{trans3D}}\, q_i^{\mathrm{vib}}\, q_i^{\mathrm{rot}}\right) \quad . \tag{2.51}$$

Assuming an ideal gas and replacing q_i^{trans3D} with the expression in Eq. (2.46), divided by N_i particles within the entire volume, leads to

$$\mu_i = E_{\mathrm{gas},i} - k_B T \ln\left(q_i^{\mathrm{vib}}\, q_i^{\mathrm{rot}}\, q_i^{\mathrm{trans2D}}\, \frac{\sqrt{2\pi m_i k_B T}}{h} \frac{k_B T}{p_i A_{\mathrm{uc}}}\right) \quad . \tag{2.52}$$

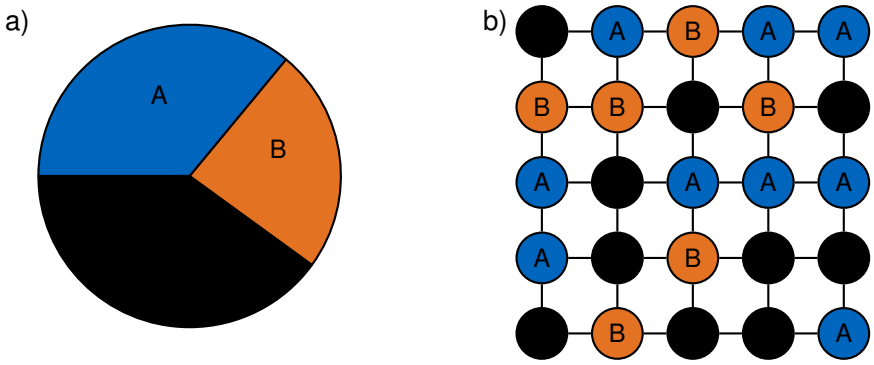


Figure 2.4: (a) Pie chart of the coverages θ of species A (blue), species B (orange), and empty sites (black) of an MFA model with a single active site type. (b) Occupation of a discrete KMC lattice with different species and empty sites using the same color coding and average coverages as in (a).

Inserting the difference between the adsorbed state in Eq. (2.50) and the chemical potential in Eq. (2.52) in Eqs. (2.48) and (2.49) for an activated adsorption process yields

$$k_{i,s}^{\text{ad,act}} = \frac{A_s}{A_{\text{uc}}} \frac{k_B T}{h} \exp\left(-\frac{G_a^+}{k_B T}\right) \quad . \tag{2.53}$$

In case of a non-activated adsorption the TS is assumed to be an early 2D gas-like TS (cf. orange line in Fig. 2.3). Consequently, the partition function in Eq. (2.49) of the TS is equal to the one of the IS and $E_a^+ = 0$ leading to

$$k_{i,s}^{\text{ad,non-act}} = \frac{p_i A_s}{\sqrt{2\pi m_i k_B T}} \quad . \tag{2.54}$$

Surface-bound processes, like reactions and diffusion, are derived analogously using Eq. (2.41) with $f_{\text{IS}\to\text{FS}}^{\text{TST}} = \frac{q_{\text{TS}}^{\text{vib}}}{q_i^{\text{vib}}}$ leading to

$$k_{i,s}^{\text{bound}} = \frac{k_B T}{h} \exp\left(-\frac{G_a^+}{k_B T}\right) \quad . \tag{2.55}$$

The rate constants for the reversed reactions are determined using Eq. (2.43). Finally, this leads to a full set of rate constants for a reaction network. Note that typical errors of DFT originating from an incomplete basis or an approximate functional, or due to free energy extrapolations are of the order of 0.2-0.3 eV. As the energetic barriers enter the exponent, this can lead to errors of two or three orders of magnitude in the rate constants depending on the temperature at play.[11, 15, 89, 117, 132, 137, 141, 142]

2.3.5 Mean Field Approximation

The MFA is one method to solve the master equation (Eq. (2.40)). Within this approach the occupancy of any surface site type s with an adsorbate i is abstracted to an average

coverage $\theta_{i,s}$. This assumption entails a random distribution of all species and the loss of a spatially resolved picture (cf. Fig. 2.4a). Consequently, diffusion processes are infinitely fast and no lateral interactions between adsorbates are considered but only indirect interactions due to blocking of a specific site type.[19, 118]

Putting this into a mathematical form leads to the rate expression

$$r_{vw}(t) = N_{vw} k_{vw} \prod_{s \in v} \theta_{i,s}(t) \quad , \tag{2.56}$$

with the site connectivity N_{vw} between the sites in basin v and w and the coverage $\theta_{i,s}(t)$ of species i on site s at time t. The reaction network of Eqs. (2.39a) to (2.39c) with only one site type is then a set of coupled differential equations:

$$\frac{d}{dt}\theta_\mathrm{A}(t) = k_\mathrm{A}^+ \theta_*(t) - k_\mathrm{A}^- \theta_\mathrm{A}(t) - k_\mathrm{C} \theta_\mathrm{A}(t) \theta_\mathrm{B}(t) \quad , \tag{2.57a}$$

$$\frac{d}{dt}\theta_\mathrm{B}(t) = k_\mathrm{B}^+ \theta_*(t) - k_\mathrm{B}^- \theta_\mathrm{B}(t) - k_\mathrm{C} \theta_\mathrm{A}(t) \theta_\mathrm{B}(t) \quad \text{, and} \tag{2.57b}$$

$$\frac{d}{dt}\theta_\mathrm{C}(t) = k_\mathrm{C} \theta_\mathrm{A}(t) \theta_\mathrm{B}(t) \quad , \tag{2.57c}$$

with the only inputs being the rate constants $k_i^\pm$ and the constraint $\sum_i \theta_i = 1$. These equations are solved toward a steady state solution which leads to the coverages θ_i and consequently a model agnostic to the actual spatial composition. Within this approximation the adsorbed species do not interact with each other. However, it is known that such interactions might play a decisive role and depending on the nature of the different interactions ordered patterns might be formed on the surface, thereby, influencing the interplay between the explicit species and possibly leading to a break down of the MFA.[19, 118, 143]

2.3.6 Kinetic Monte Carlo

KMC is another method to solve the master equation. In contrast to the MFA, it is a numerical approach, which allows to handle the curse of dimensionality to reach an arbitrarily accurate solution by describing the system as a Markov chain of independent events, i.e., the system loses its memory about the preceding events.[139] These events are the possible elementary processes on the basis of the current lattice configuration and represent discrete transitions from one lattice configuration to another. One possible lattice configuration for the two species A (blue) and B (orange) of the elementary steps in Eqs. (2.39a) to (2.39c) is shown in Fig. 2.4b, which is one of several configurations of the mean coverage shown in Fig. 2.4a. To obtain the probability $P_v(t)$ of a system to be in state v at time t within the master equation (Eq. (2.40)), an ensemble of generated stochastic trajectories or – if the system is in steady state and ergodic – a sufficiently long trajectory is averaged.

Variable Step-Size Method

The variable step-size method (VSSM) is one approach for the implementation of KMC, which belongs to the rejection-free algorithms (cf. Fig. 2.5 blue background).[144–146] Compared to rejection approaches for which execution of a process might be rejected, the

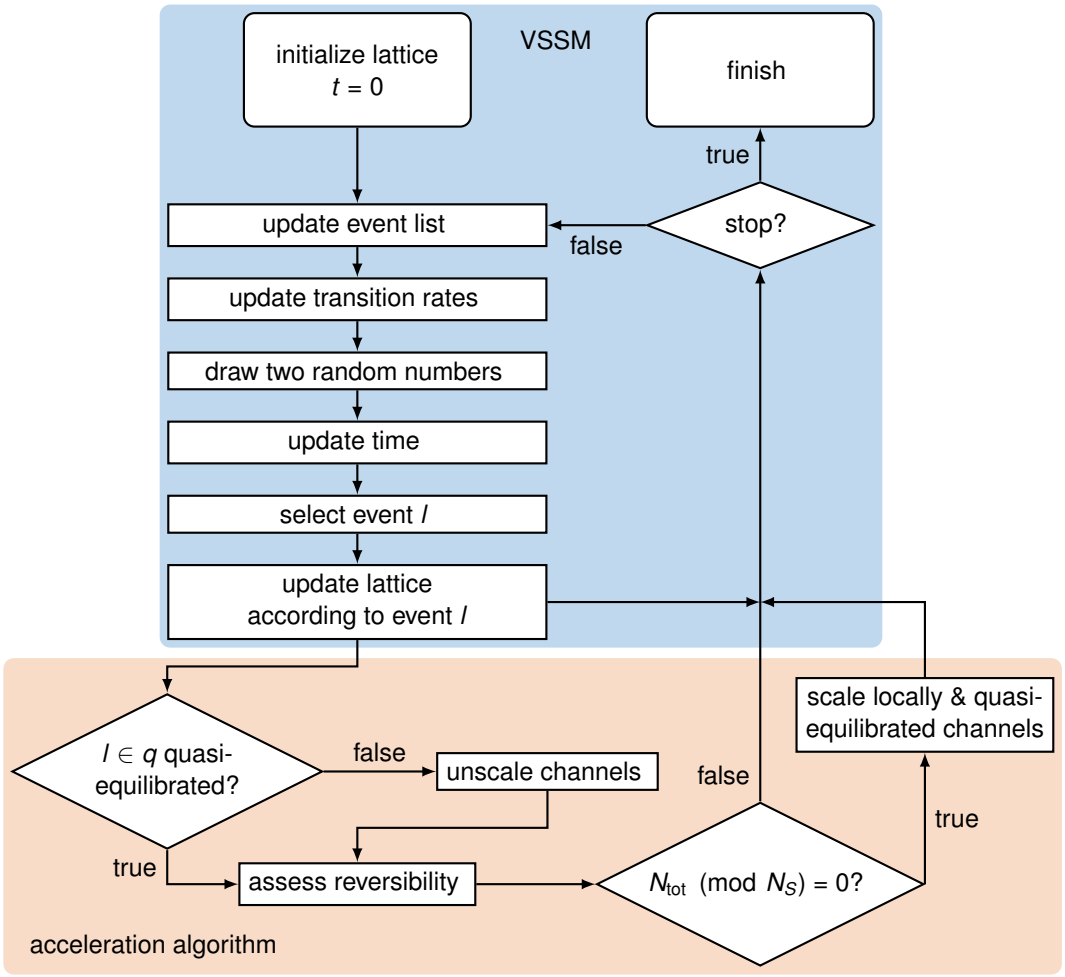


Figure 2.5: Flow chart of the VSSM KMC algorithm (blue) and the additional acceleration steps (orange) with the corresponding variables and parameters as described in the text.

rejection-free algorithm always executes an event and the lattice needs to be updated in each iteration. In this algorithm, the event list is updated at each iteration together with the associated rate constants.[19, 82, 147]

The starting point of a VSSM KMC simulation can be any lattice configuration v. At first, all possible events M that can occur from v to lattice configuration w are detected and the total rate constant

$$k_{\text{tot}} = \sum_{w=1}^{M} k_{vw} \tag{2.58}$$

for which the transition probabilities are given by the rate constants k_{vw} is calculated. Then one event l leading to lattice configuration n is selected by generating a uniformly distributed

random number $\rho_1 \in]0, 1]$ fulfilling

$$\sum_{w=1}^{n} k_{vw} \geq \rho_1 k_{\text{tot}} \geq \sum_{w=1}^{n-1} k_{vw} \quad . \tag{2.59}$$

The probability distribution of such a process within time Δt to go from the PES basin v into basin w decays exponentially and follows a Poisson distribution

$$p_{vw}(\Delta t_{vw}) = k_{vw} \exp(-k_{vw} \Delta t_{vw}) \quad , \tag{2.60}$$

which is centered around the average escape time $\overline{\Delta t}_{vw} = k_{vw}^{-1}$. The probability of first escape from basin v to any other state is described by

$$p_{\text{escape}}(\Delta t) = k_{\text{tot}} \exp(-k_{\text{tot}} \Delta t) \tag{2.61}$$

and the corresponding escape time Δt for advancing the simulation $t \rightarrow t + \Delta t$ needs to be weighted by the probability distribution $p_{\text{escape}}(\Delta t)$ to arrive at

$$\Delta t = -\frac{\ln(\rho_2)}{k_{\text{tot}}} \quad , \tag{2.62}$$

with a second uniformly distributed random number $\rho_2 \in]0, 1]$.[19, 82, 147]

Acceleration Algorithm

The drawback of KMC is a consequence of its statistical nature, which requires a sufficient sampling within simulations to obtain statistically significant results. Compared to the MFA, this leads to larger computational requirements with growing reaction networks. One of the culprits responsible for these performance issues is the time-scale disparity problem, which is the discrepancy between fast and slow processes.[23] Due to the exponential dependence of the rate constants on the activation barriers, processes like diffusion and adsorption, which often have low such barriers, occur at vastly different time scales than surface bound processes with larger barriers (cf. Eq. (2.41)). Consequently, the fast reactions are sampled much more frequently (cf. Eq. (2.59)) than slow processes, which requires a longer overall run time for a sufficient sampling. Artificially raising certain barriers of fast processes or a grouping into slow and fast processes previous to the simulation mitigates the problem, but care has to be taken to maintain ergodicity and still achieve equilibration of the adlayer ordering between the slower processes.[23, 135, 136, 148–150] Other methods to speed up KMC simulations also involve parallelization of a KMC algorithm, which is, however, not subject of this work.[23, 91]

The extended phenomenological kinetics (XPK) method couples KMC for equilibrating diffusion and the MFA for evolving reaction processes to avoid the time-scale disparity problem. This approach may, however, influence the kinetic behaviour within the simulation as diffusion processes are always equilibrated in the MFA step and resulting diffusion limitations cannot be captured.[3, 16, 118, 150, 151] A different approach to speed up KMC simulations was introduced by Chatterjee and Voter [22] and adapted by Dybeck et al.[23] It uses

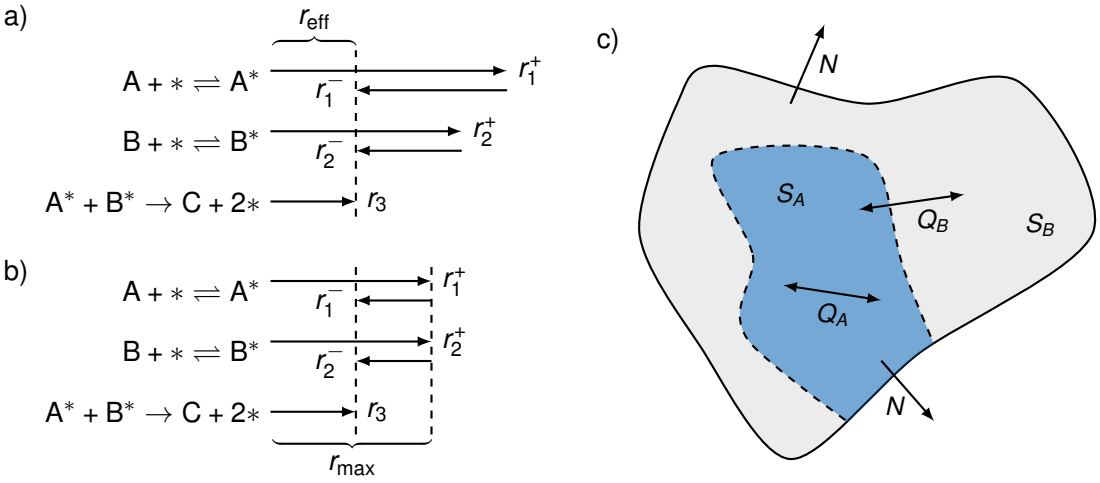


Figure 2.6: (a) Unscaled and (b) scaled hypothetical rates $r_q^{\pm}$ of the Langmuir-Hinshelwood mechanism shown in Eqs. (2.39a) to (2.39c). Each step (forward and backward event) represents a reaction channel. (c) Superbasin S (all quasi-equilibrated processes, see Eq. (2.63)), which is the union of the blue region S_A of locally equilibrated channels ($n_c^+ + n_c^- \geq n_e$) and the gray region S_B of not locally equilibrated channels. Quasi-equilibrated channels Q_A stay within S_A, while locally equilibrated channels Q_B explore new regions in S_B within the superbasin S. Non-equilibrated channels N lead to the exploration of a new superbasin. Adapted with permission from *J. Chem. Theory Comput.* 2017, 13, 1525–1538. © 2017 American Chemical Society.

reaction channels and superbasins to achieve a temporal acceleration of a KMC simulation without the need to resort to the MFA or prior knowledge via scaling of fast processes while not affecting slower processes. This concept is shown in Figs. 2.6a and 2.6b as well as Fig. 2.6c. The respective additional steps for a KMC simulation are shown in Fig. 2.5 (orange background). A reaction channel q entails the forward and backward direction of a process to maintain detailed balance. For the case of the Langmuir-Hinshelwood mechanism (Eqs. (2.39a) to (2.39c)) in Fig. 2.6a, each of the three channels q is represented by their respective rates $r_q^{\pm}$. The two fast processes (Eqs. (2.39a) and (2.39b)) are scaled by the algorithm to $r_{\text{max}} = N_f r_{\text{eff}}$, where the buffer parameter N_f is a measure of how often on average the forward and backward events are sampled before the slow process (r_3, Eq. (2.39c)) is executed. In the following the mathematical derivation of the algorithm is described. The starting point is to partition the individual channels according to the reversibility condition

$$\frac{|n_q^+ - n_q^-|}{n_e} \leq \delta \tag{2.63}$$

into quasi-equilibrated (Q) and non-equilibrated (N) channels depending on whether the inequality is fulfilled or not. Here, n_q^+ and n_q^- are the counts of forward and backward exe-

cutions of channel q over the last n_e executions, which is an adjustable parameter, and δ is the threshold parameter. A superbasin S is defined by all lattice states reachable via quasi-equilibrated processes Q. Starting a simulation with not locally equilibrated channels, which means that $n_q^+ + n_q^- < n_e$, each channel is subsequently classified according to Eq. (2.63). Fast channels will be classified as quasi-equilibrated and subject to scaling according to

$$k_l' = \alpha_q k_l \quad , \tag{2.64}$$

with the scaled (k_l') and non-scaled rate constant (k_l) of event $l \in q$ as well as the scaling factor $\alpha_q \in [0, 1]$. The scaling is performed in intervals of N_S of the total simulation steps N_{tot}, where α_q is evaluated via

$$\alpha_q = N_f \frac{2r_S}{r_{q,S}} \quad , \tag{2.65}$$

with the buffer parameter N_f and the effective superbasin escape rate

$$r_S = \sum_{q \in N, Q_B} r_{q,S} \quad . \tag{2.66}$$

N is the set of non-equilibrated channels and Q_B the set of quasi-equilibrated but not locally equilibrated channels, which take the system from the explored region S_A in the unexplored region S_B of the superbasin S. The reaction channels Q_A in Fig. 2.6c on the other hand resemble quasi-equilibrated and locally equilibrated channels, which are confined to the explored region S_A. The average rate $r_{q,S}$ of reaction channel q is determined via

$$r_{q,S} = \frac{1}{\sum_{n \in S} \Delta t_n} \sum_{n \in S} k_q(n) \Delta t_n \quad , \tag{2.67}$$

with the sum going over all steps n performed in superbasin S and the time Δt_n of each step. The rate of channel q is obtained via the sum over all rate constants k_l of event $l \in q$ that can occur at step n:

$$k_q(n) = \sum_{l \in q} k_l(n) \quad . \tag{2.68}$$

If the chosen event is non-equilibrated ($l \in N$), the network is assumed to have left a superbasin, which implies that a slow process was executed. Consequently, all scaling factors α_q are reset to unity, all execution counts n_e are reset to zero, i.e., all channels are not locally equilibrated, and the partitioning into quasi- and non-equilibrated persists within the new superbasin. From this point on the algorithm re-enters the standard VSSM loop.[22, 23, 137]

2.3.7 Degree of Rate Control

The interplay of adsorption and desorption processes of atoms or molecules onto the surface as well as surface bound processes of adsorbed species forms the reaction network,

which contains all possible pathways from one state to another. In order to analyze the peculiarities of such a reaction network, the limiting factors within the different possible reaction mechanisms can be determined. As described above, the rate of a certain transformation is based on the energetics of the involved species as activated steps going from the IS to the FS require to overcome a TS, which determines the energy barriers for the forward and backward process. Usually only one or a small subset of elementary steps limit the overall transformation and impact the product selectivities. These steps are the (partly) RDSs. A convenient tool to identify RDSs is sensitivity analysis, which relates the change of the overall turnover frequency (TOF) of the formation of product i with a change in the speed of a single elementary reaction resembled by the free energy of formation G_j of a TS j:[152, 153]

$$X_{ij} = \left(\frac{\partial \ln (r_i)}{\partial \left(-G_j / k_B T \right)} \right)_{G_{j \neq o}}, \tag{2.69}$$

with the degree of rate control (DRC) matrix X_{ij}, the production (consumption) rate r_i of product (reactant) i, the Boltzmann constant k_B, and the temperature T. All o other formation energies $G_{j \neq o}$ are kept constant leading to a measure of the influence of a single elementary reaction step on the TOF. If the DRC is positive (negative), the TOF is increased by a more strongly (weakly) bound species j. Consequently, the network needs to contain the correct elementary reactions and accurate energetics for these steps as the remaining steps are unimportant once the RDSs are identified for a certain set of reaction conditions. This renders sensitivity analysis also usually much more robust compared to DFT due to its associated errors. However, one has to keep in mind that this concept neglects the interdependence among adsorption energetics exploited, e.g., by scaling relations. In essence the virtue of screening studies is that relative activities are much more meaningful than single absolute TOFs due to the correlations inherent in the energetics predicted by DFT functionals, which in most cases lead to error cancellation. Finally, the success of such studies predicting the well known activity volcanos rely on relative activities. However, it is important to point out that the common approach is to use a predefined reaction network – based on the insights obtained from catalysts around the top of the volcano – for all transition metals (TMs) under consideration, which might not be a valid assumption.[133, 152–157]

2.3.8 Comparison Between the MFA and KMC

The averaging of the surface coverage within the MFA simulations leads to the problem that the individual interactions between adsorbates can not be accounted for. As described in Sec. 2.3.5, the only possibility to include adsorbate correlations is to introduce coverage-dependent adsorption energies, which only affect the respective average coverages. KMC studies, however, have shown that differences in the local environment due to spatial correlations can have a significant influence on the catalytic predictions.[3, 16, 118]

Especially second-order processes are associated with the differences in the simulation results. Based on a KMC simulation the pair probability $N_2(\mathrm{A}, \mathrm{B})$ resembles the probability

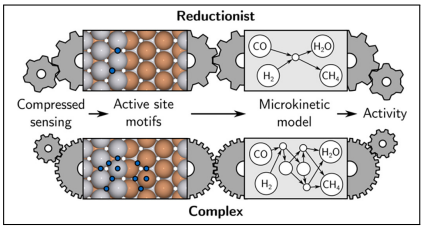
of two reacting species A and B residing in a state where a reaction is possible. It is obtained by counting these states over the whole simulation and normalizing by the simulation time and the system size. On the other hand, based on the average coverage of these two species θ_A and θ_B within the KMC simulation and the site connectivity $N_{A,B}$, the MFA-assumed probability is given by $N_{A,B}\theta_A(t)\theta_B(t)$ (cf. Eq. (2.56)). In case the ratio of the pair probability divided by the MFA-assumed probability

$$\frac{N_2(A,B)}{N_{A,B}\theta_A(t)\theta_B(t)} \tag{2.70}$$

deviates substantially from unity, the assumptions within the MFA lead to errors and it breaks down. A value lower than one denotes that the reaction is less likely to occur in KMC, while a larger value denotes the opposite. This can be translated into a lower (higher) configurational entropy in the MFA-assumed well-mixed state, which in turn influences the free energy landscape.[3, 137]

3 Publications

3.1 Active Site Representation in First-Principles Microkinetic Models: Data-Enhanced Computational Screening for Improved Methanation Catalysts



Martin Deimel, Karsten Reuter, and Mie Andersen
ACS Catal. **10**, 13729 (2020).
DOI: 10.1021/acscatal.0c04045

Summary

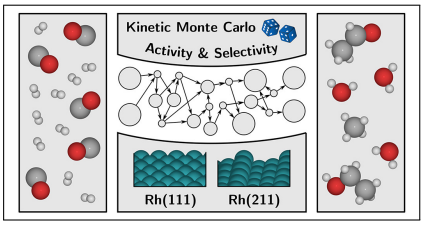
The goal of this work is to investigate the effect of the active site representation of potential transition metal (TM) catalyst surfaces for the carbon monoxide (CO) methanation reaction. With the help of machine learning (ML) techniques, we overcome the bottleneck of electronic structure theory, which is especially crucial for computational screening studies. The catalyst surfaces used within the article entailed pure elements as well as binary alloys, which offer – due to their different atom composition – a larger variety of diverse active site patterns. This is resembled by a larger spectrum of different adsorption energies available for the reaction intermediates and the potential for a benign catalytic pathway. To simulate the effect of such an active site resolution the activity predictions from three mean-field microkinetic models of different complexity were compared and differences and similarities analyzed. Based on a Sure Independence Screening and Sparsifying Operator (SISSO) descriptor, which was fit to density functional theory (DFT) data and was based only on properties of the clean surface, the required energetic input of the whole surface could be obtained with a single DFT calculation per catalyst material.

The insight obtained from this project is that a more reductionist active site representation has a tendency to overestimate the activity of pure TMs compared to a more detailed model while the opposite is the case for layered bimetallic alloys. This discrepancy is traced back to shortcomings in an adequate representation of the variety of accessible active sites and energetic ordering for the adsorbates involved in the rate-determining step (RDS) within the simpler models. For most of the potential catalyst surfaces the RDS is the H-assisted CO dissociation. In particular a weak interaction of the surface with the emerging OH adsorbate within this step on layered alloys causes a smaller activation barrier and a larger turnover frequency (TOF), which can not be resolved with a coarser active site representation and would be missed. Using ML for the fast and inexpensive prediction of adsorption energies opens up the possibility to use detailed microkinetic models, which require a much larger number of input energetics, for large screening studies and gain a better mechanistic understanding of the catalytic process.

Individual Contributions

Following the idea of Mie Andersen, I extended the DFT calculated data set published in Ref. 19 by the adsorption energetics on the high symmetry sites and the respective primary features of the adsorbates CH_2 and CH_3 and the TMs Co and Fe. The original database entailed the energetics of H, C, CH, CO, O, and OH on the face-centered cubic (fcc) (211) facets of pure Ni, Cu, Ru, Rh, Pd, Ag, Ir, Pt, and Au as well as the four single-atom alloys Ag@Cu, Pt@Rh, Pd@Ir, and Au@Ni, and the four AB bimetallic alloys AgPd, IrRu, PtRh, and AgAu. The two additional TMs and species are relevant catalyst materials and intermediates for the methanation reaction. As the most stable phase of Fe is a body-centered cubic (bcc) crystal, I determined the adsorption energies on the bcc(210) facet and, similarly, calculated in addition to the fcc(211) facet of Co, the energetics for a stepped hexagonal close packed (hcp) (0001) facet. On the basis of this data set I fit and evaluated a SISSO [103, 104] descriptor based on a training and validation set, which consisted of a random selection of 80% and 20%, respectively, of the available data of each individual adsorbate. I then used this descriptor to predict the adsorption energies of a selection of binary alloy surfaces of all the mentioned TMs within the space groups 164 and 187 (hcp-like) as well as 221 and 225 (fcc-like) obtained by careful selection of bulk phases from Automatic Flow (AFLOW)-convex hull (CHULL) database searches to model their catalytic activity. To this end, simulations of three mean-field microkinetic models with an increased active site resolution were conducted and compared by me. The most detailed model was taken from Ref. [137], the intermediate one from Ref. [71], while the reductionist model was introduced in this work. The manuscript was collectively written and edited by all authors.

3.2 Selectivity Trends and Role of Adsorbate–Adsorbate Interactions in CO Hydrogenation on Rhodium Catalysts



Martin Deimel, Hector Prats, Michael Seibt, Karsten Reuter, and Mie Andersen
ACS Catal. **12**, 7907 (2022).
DOI: 10.1021/acscatal.2c02353

Summary

This project on the carbon monoxide (CO) hydrogenation reaction on the Rh(111) and Rh(211) facets investigates the activity and selectivity trends employing accelerated first-principles kinetic Monte Carlo (KMC) simulations. The focus lies on the influence of lateral interactions parametrized from a cluster expansion (CE) model and a comparison of our results with and without lateral interactions to the theoretical work on these surfaces by Yang et al.[21] using a mean field approximation (MFA) approach with coverage dependent reaction rates as well as the experimental results of Schumann et al.[51] for Rh nanoparticles of different sizes. The acceleration algorithm is crucial due to the large reaction networks for both our Rh(211) model, which consists of 34 reaction and 19 diffusion steps, and the Rh(111) model with 16 reaction and 5 diffusion steps, and the concomitant time-scale disparity problem usually occurring due to fast diffusion and slow reaction steps described in 2.3.6. On Rh(111) it can be seen from the MFA data that lateral interactions have a huge impact on the activity and selectivity. We rationalize this by the differences in the coverages. The secondary reaction step of CHOH dissociation – the rate-limiting step (RLS) on the Rh(111) facet – requires a free neighboring terrace site, which is more probable when the repulsive CO-CO interactions cause a lower coverage of the dominant CO adsorbate. A comparison between KMC and the MFA including lateral interactions reveals that the selectivity differs below $\sim$580 K. The MFA predicts selectivity toward acetaldehyde while KMC predicts methane (CH_4) selectivity in agreement with the experimental findings on large nanoparticles. On the other hand, there is only a small influence on the activity on Rh(211) with and without lateral interactions. The coverages are less affected by the interactions and the effective barrier of the RLS, which is the water (H_2O) formation at the terrace site, remains similar. From an analysis of the pair probabilities without interactions, localized lattice inhomogeneities are identified, which can only be seen within KMC due to the average coverage used in the MFA. These cause lower turnover frequencies (TOFs) and different RLSs in the KMC than the MFA model. In an attempt to identify the reason for the differences of the KMC activities and selectivities to the experimental results, we found that a correction to the common CO overbinding of generalized gradient approximation (GGA) functionals leads to an overall better agreement. This can be traced back to the larger influence on the acetaldehyde species – containing two C atoms – as two individual CO species are involved in the formation as opposed to CH_4, where only one CO is involved.

Individual Contributions

Based on the initial idea of Mie Andersen of combining the acceleration algorithm introduced by Dybeck et al.[23] for KMC simulations with the on-the-fly calculation of rate constants for the treatment of lateral interactions, the project resulted from the work by Michael Seibt,[150] who implemented and tested the algorithm in *kmos*.[147] Michael Seibt implemented the Bayesian error analysis and with Hector Prats jointly developed the KMC reaction networks for the Rh(111) and Rh(211) facets based on the networks and energetics by Yang et al.[21] Some additional calculations for intermediates and transition states (TSs) as well as for the CE were required for these networks, which have been performed by Hector Prats. I updated and adapted these networks. All KMC simulations were performed by me including a sensitivity analysis, calculations for the impact of errors in the calculated CO adsorption energy, as well as the determination of the KMC and MFA-assumed pair probabilities. I used the Rh(111) and adapted the Rh(211) MFA reaction networks by Yang et al.[21] to match the KMC networks in order to be able to compare the results from my Catalysis Microkinetic Analysis Package (CatMAP) [143] simulations and trace back differences solely to the different approaches and approximations. The manuscript was collectively written and edited by Mie Andersen, Hector Prats, Karsten Reuter, and myself.

3.3 Further Work

Additionally to the publications above two articles have been published during the course of my time at the Chair for Theoretical Chemistry and Catalysis Research Center. These articles are not directly related to the topic of this thesis and therefore are solely mentioned for the sake of completeness.

- **Interface between graphene and liquid Cu from molecular dynamics simulations**
  Juan Santiago Cingolani, Martin Deimel, Simone Köcher, Christoph Scheurer, Karsten Reuter, and Mie Andersen
  J. Chem. Phys. **153**, 074702 (2020).
  DOI: 10.1063/5.0020126

- **Generalized molecular solvation in non-aqueous solutions by a single parameter implicit solvation scheme**
  Christoph Hille*, Stefan Ringe*, Martin Deimel, Christian Kunkel, William E. Acree, Karsten Reuter, and Harald Oberhofer
  J. Chem. Phys. **150**, 041710 (2019).
  DOI: 10.1063/1.5050938
  * These authors contributed equally to the work.

4 Summary, Conclusions, and Outlook

Microkinetic modeling (MKM) advanced to a very important tool to aid in understanding heterogeneous catalytic processes and promoting catalysis research.[15, 19, 82] The underlying high-dimensional potential energy surface (PES) forms the basis to such modeling, as it describes the entire system and the transitions between different states on a catalyst surface. Using first-principles (1p) quantum mechanical methods such as density functional theory (DFT) to obtain single points on the PES allow to describe the system in terms of rare events, the so-called elementary reactions. These points represent minima (adsorbed species) and saddle points (transition states (TSs)) at and between specific active sites. In combination with the mean field approximation (MFA), this forms the basis of catalyst screening studies. Such studies take advantage of methods to approximate the PESs of possible catalyst materials from the information on other PESs. These methods include linear scaling relations and more advanced machine learning (ML) methods.[19] Despite the approximations included in such theoretical studies, the overall trend is captured quite well.[21] The MFA approach proved to be very helpful in explaining experimental observations and such calculations could guide further towards promising catalyst materials at a fraction of the cost associated with experimental setups. Refining the models itself as well as the approximations for obtaining the data will naturally lead to a better description of the processes under investigation. However, the assumption of a mean coverage without any interactions between adsorbates leads to problems especially when in regions with the highest activity. Realizing that the surface within these regions is simultaneously covered with different species and exhibiting very different local environments, it turned out that such a neglect of interactions and a site resolved picture is too crude for an atomic-scale understanding of the important surface processes. In this regard, kinetic Monte Carlo (KMC) in combination with a cluster expansion (CE) approach to model lateral interactions between adsorbed species is the method of choice for in-depth mechanistic insight at the cost of more intricate calculations.

In our investigations on the methanation reaction on binary alloy catalysts we realized that the choice of active sites has a decisive impact on the simulation results.[2] It is essential to include all high symmetry sites to obtain an accurate description of the catalytic process. This was illustrated by comparing the catalytic activities obtained from different MFA models with increasing active site resolutions. As a key aspect for the determination of novel catalyst materials with improved performance is efficiency, we predicted the required extensive

energetic data with the ML method Sure Independence Screening and Sparsifying Operator (SISSO). While this approach provides the necessary performance and narrows down the space of possible catalyst materials to just a few promising ones, the results preclude insight on the atomic-scale, which contains possibly decisive factors. In this regard, the KMC study of the carbon monoxide (CO) hydrogenation reaction toward C_{2+} oxygenates revealed such factors. Overcoming the time-scale disparity problem using the acceleration algorithm by Dybeck et al.[23] together with on-the-fly (otf) rate constants within *kmos* [147] allows to reach out to such large reaction networks. The local coverages impacting initial states (ISs), TSs, and final states (FSs) via different lateral interactions can influence the energetics and limit diffusion. Consequently, reaction channels are enabled or disabled, leading to states with immobile adsorbates. The detailed understanding gained about the micro-scale factors influencing selectivity and activity, which is mostly not possible in an experimental setup, sets the stage for targeted catalyst design and upscaling approaches.

As insightful the results of the two publications are, there are also some issues to be addressed in future research. While it is possible to tackle large and complex systems, it is important to keep in mind that the quality of the MKM simulations severely depend on the accuracy of the input energetics due to the exponential dependency of the rate constants on the barriers. The prominent failure of generalized gradient approximation (GGA) functionals in accurately calculating CO adsorption [57–59] is highlighted in Ref. 3. As a result, the calculated activities and selectivities can vary greatly. To this end, ML approaches for the prediction of the energetic input are a suitable means for even more reliable energetics without largely impacting computational demands.[158, 159] Moreover, ML can also be used to obtain possible catalyst materials in a bottom-up procedure by mapping ideal parameters to a structure.[94] Another, yet already tackled topic is the appearance of multidentate species, especially when dealing with larger adsorbates. The established approach uses graph theory to treat these cases. Such graph-theoretical ideas are also used to model a system consisting of different lattices like nanoparticles.[89] While these concepts have already been applied, *kmos* currently does not implement them, which might be a valuable contribution in the future. Lastly, topological information of real catalytic systems at *operando* conditions is scarce. Consequently, and due to the sheer number of possibilities, theoretical studies focus on idealized facets, which neglect the influence of dynamical phenomena and phase transitions. All in all, it will be very interesting to see how the field of MKM will develop with these and other daunting challenges in the years to come.

Acknowledgments

First of all, I want to thank Karsten Reuter for giving me the opportunity and support to write this PhD thesis. Thank you for letting me follow my aspirations to enter the field of theoretical and computational chemistry by welcoming me to the group already back in 2014 for my bachelor's thesis, and keeping me around for all these years.

Equally, I would like to thank Mie Andersen for her continuous scientific and non-scientific support during the time of my PhD and for supervising, advising, and taking care of me during the development of this thesis.

Further thanks go to Ruth Mösch for all the nice conversations and the bureaucratic and administrative help throughout the years. She always had an open ear and valiantly held the line with me in Munich, while the rest of the group gradually migrated to new endeavours in Berlin. I want to thank Harald Oberhofer for supervising my bachelor's thesis and always having a minute for nerdy conversations. Moreover, I want to thank all the people that kept our IT infrastructure alive for all these years. While it may have caused some sleepless nights and frustration, I really enjoyed this duty and especially the incredible team behind it. Thank you Christoph Scheurer, for always sharing your tremendous knowledge, Georg Michelitsch, for your support and introducing me to the team, Matthias Kick, for taking care of the software side and reminding me that most of these things do not need to be taken care of instantly, David Egger, for the best instructor and hardware team member, Simon Wengert, for his software support duty as well as his calm nature, Christoph Muschielok, for your help and support especially towards the end, as well as Simeon Beinlich and Christian Kunkel, for taking care of the first level support. Moreover, I want to thank my office mates Santiago Cingolani and Simon Wengert as well as almost office mate Sina Stocker for the priceless conversations, Thursday evenings, and all the other times we spent together. Thank you people of the last batch staying or remaining in Munich; among them Cristina Grosu, Christoph Muschielok, Ruth Mösch, Harald Oberhofer, Sina Stegmaier, and Jakob Timmermann. I would also like to thank Sebastian Matera for proofreading of the thesis as well as Hector Prats and Michael Seibt for their work and help on our article. From the new batch Younes Hassani Abdollahi, Sina Dortaj, Steffen Kangowski, and Julia Pach should not be left unmentioned. Furthermore, I would like to express my fullest gratitude to all other former members of the group at the TU Munich as well as the "new old" and new ones at the Fritz Haber Institute in Berlin for the numerous talks, discussions, coffee breaks, conferences,

workshops, and experiences during the last years. To name a few of them as there are just too many: Ahmad Agung, Nicolas Bergmann, Albert Bruix, Vanessa Bukas, Ke Chen, Frederic Felsen, Simiam Ghan, Patrick Gütlein, Hendrik Heenen, Simone Köcher, Elisabetta Landini, Yonghyuk Lee, Haobo Li, Juan Lorenzi, Arobendo Mondal, Daniel Opalka, Chiara Panosetti, Craig Plaisance, Christoph Schober, Markus Sinstein, Carsten Staacke, Thomas Stecher, Hanna Türk, and Wenbin Xu. With a lot of these people I became good friends and this, I guess, is what a good working environment is about.

Moreover, I want to thank the people and friends, who I met during and who accompanied me through the chemistry studies, Andreas, Benedikt, Carlos, Denis, Ines, Joachim, Jonas, Julia, Lorenz, Marco, Max, Sophie, Tassilo, Thomas, and many, many more, for the precious times and memories.

I gratefully acknowledge funding from the LMCat Project of the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 736299, the Solar Technologies Go Hybrid Initiative of the State of Bavaria, and the E-Conversion Cluster of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2089/1-390776260. Furthermore, I gratefully acknowledge support from the Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, and the Jülich Supercomputing Centre.

Finally, I want to thank all of my close friends – you know who you are – and above all my family. You are the ones, who were guiding and supporting me through the good and the bad times and enabling me to follow through with my path.

Munich, November 2022

Bibliography

[1] J. Marinić, *Mai Thi Nguyen-Kim im Interview: "Wissenschaft ist keine Glaskugel"*, Accessed: 2022-10-01, Hessischer Rundfunk, Oct. 2021 (cit. on p. III).

[2] M. Deimel, K. Reuter, and M. Andersen, ACS Catal. **10**, 13729 (2020) (cit. on pp. i, 2, 6, 14–18, 39).

[3] M. Deimel, H. Prats, M. Seibt, K. Reuter, and M. Andersen, ACS Catal. **12**, 7907 (2022) (cit. on pp. i, 2, 3, 7, 11, 27, 30, 31, 40).

[4] J. W. von Goethe, J. W. Döbereiner, and J. Schiff, *Briefwechsel zwischen Goethe und Johann Wolfgang Döbereiner, 1810–1830. Herausgegeben und erläutert von Julius Schiff* (Hermann Böhlaus Nachfolger, 1914) (cit. on p. 1).

[5] G. Ertl, Angew. Chem. Int. Ed. **47**, 3524 (2008) (cit. on p. 1).

[6] M. Appl, "Ammonia, 1. Introduction," in *Ullmann's Encyclopedia of Industrial Chemistry* (John Wiley & Sons, Ltd., 2011) (cit. on p. 1).

[7] F. Fischer, H. Tropsch, and P. Dilthey, Brennst.-Chem. **6**, 265 (1925) (cit. on p. 1).

[8] K. Cheng, J. Kang, D. L. King, V. Subramanian, C. Zhou, Q. Zhang, and Y. Wang, in *Advances in catalysis*, Vol. 60, edited by C. Song (Academic Press, 2017), pp. 125–208 (cit. on pp. 1, 5, 6).

[9] J. J. Bravo-Suárez, R. V. Chaudhari, and B. Subramaniam, "Design of Heterogeneous Catalysts for Fuels and Chemicals Processing: An Overview," in *Novel Materials for Catalysis and Fuels Processing*, edited by J. J. Bravo-Suárez, M. K. Kidder, and V. Schwartz (American Chemical Society, 2013) Chap. 1, pp. 3–68 (cit. on p. 1).

[10] K. S. Egorova and V. P. Ananikov, Organometallics **36**, 4071 (2017) (cit. on p. 1).

[11] I. Chorkendorff and J. W. Niemantsverdriet, *Concepts of Modern Catalysis and Kinetics*, Third ed. (Wiley-VCH, 2017) (cit. on pp. 1, 11, 21, 22, 24).

[12] G. Ertl, Angew. Chem. Int. Ed. **48**, 6600 (2009) (cit. on p. 1).

[13] K. Rajan, Annu. Rev. Mater. Res. **38**, 299 (2008) (cit. on p. 1).

[14] C. A. Gaggioli, S. J. Stoneburner, C. J. Cramer, and L. Gagliardi, ACS Catal. **9**, 8481 (2019) (cit. on p. 1).

[15] A. H. Motagamwala and J. A. Dumesic, Chem. Rev. **121**, 1049 (2021) (cit. on pp. 1, 7, 19–22, 24, 39).

[16] B. Temel, H. Meskine, K. Reuter, M. Scheffler, and H. Metiu, J. Chem. Phys. **126**, 204711 (2007) (cit. on pp. 1, 2, 27, 30).

[17] L. J. Broadbelt and R. Q. Snurr, Appl. Catal., A **200**, 23 (2000) (cit. on p. 2).

[18] J. K. Nørskov, F. Abild-Pedersen, F. Studt, and T. Bligaard, Proc. Natl. Acad. Sci. U.S.A. **108**, 937 (2011) (cit. on pp. 2, 10).
[19] M. Andersen, S. V. Levchenko, M. Scheffler, and K. Reuter, ACS Catal. **9**, 2752 (2019) (cit. on pp. 2, 11, 16, 19, 21, 25–27, 35, 39).
[20] M. Neurock and E. W. Hansen, Comput. Chem. Eng. **22**, S1045 (1998) (cit. on p. 2).
[21] N. Yang, A. J. Medford, X. Liu, F. Studt, T. Bligaard, S. F. Bent, and J. K. Nørskov, J. Am. Chem. Soc. **138**, 3705 (2016) (cit. on pp. 2, 6, 11, 36, 37, 39).
[22] A. Chatterjee and A. F. Voter, J. Chem. Phys. **132**, 194101 (2010) (cit. on pp. 2, 27, 29).
[23] E. C. Dybeck, C. P. Plaisance, and M. Neurock, J. Chem. Theory Comput. **13**, 1525 (2017) (cit. on pp. 2, 27, 29, 37, 40).
[24] M. Stamatakis and S. Piccinin, ACS Catal. **6**, 2105 (2016) (cit. on pp. 2, 11).
[25] A. Kaithal, M. Hölscher, and W. Leitner, Chem. Sci. **12**, 976 (2021) (cit. on p. 5).
[26] J. Kopyscinski, T. J. Schildhauer, and S. M. A. Biollaz, Fuel **89**, 1763 (2010) (cit. on p. 5).
[27] H. Hiller, R. Reimert, and H.-M. Stönner, "Gas Production, 1. Introduction," in *Ullmann's Encyclopedia of Industrial Chemistry* (John Wiley & Sons, Ltd., 2011) (cit. on p. 5).
[28] P. T. Anastas and J. C. Warner, *Green Chemistry: Theory and Practice* (Oxford University Press, 2000) (cit. on p. 5).
[29] P. W. N. M. Leeuwen and C. Claver, eds., *Rhodium Catalyzed Hydroformylation*, Catalysis by Metal Complexes (Springer Dordrecht, 2002) (cit. on p. 5).
[30] A. J. Ragauskas, C. K. Williams, B. H. Davison, G. Britovsek, J. Cairney, C. A. Eckert, W. J. Frederick, J. P. Hallett, D. J. Leak, C. L. Liotta, J. R. Mielenz, R. Murphy, R. Templer, and T. Tschaplinski, Science **311**, 484 (2006) (cit. on p. 5).
[31] E. van Steen and M. Claeys, Chem. Eng. Technol. **31**, 655 (2008) (cit. on p. 5).
[32] R. Franke, D. Selent, and A. Börner, Chem. Rev. **112**, 5675 (2012) (cit. on p. 5).
[33] B. Kumar, J. M. Smieja, A. F. Sasayama, and C. P. Kubiak, Chem. Commun. **48**, 272 (2012) (cit. on p. 5).
[34] N. Dahmen, E. Henrich, E. Dinjus, and F. Weirich, Energy Sustain. Soc. **2**, 3 (2012) (cit. on p. 5).
[35] R. C. Baliban, J. A. Elia, and C. A. Floudas, Energy Environ. Sci. **6**, 267 (2013) (cit. on p. 5).
[36] M. Bertau, H. Offermanns, L. Plass, F. Schmidt, and H.-J. Wernicke, eds., *Methanol: The Basic Chemical and Energy Feedstock of the Future* (Springer Berlin, Heidelberg, 2014) (cit. on p. 5).
[37] S. Xie, Q. Zhang, G. Liu, and Y. Wang, Chem. Commun. **52**, 35 (2016) (cit. on p. 5).
[38] S. Hernández, M. Amin Farkhondehfal, F. Sastre, M. Makkee, G. Saracco, and N. Russo, Green Chem. **19**, 2326 (2017) (cit. on p. 5).
[39] S. R. Foit, I. C. Vinke, L. G. J. de Haart, and R.-A. Eichel, Angew. Chem. Int. Ed. **56**, 5402 (2017) (cit. on p. 5).
[40] B. Cornils, A. Börner, R. Franke, B. Zhang, E. Wiebus, and K. Schmid, "Hydroformylation," in *Applied Homogeneous Catalysis with Organometallic Compounds*, edited

by B. Cornils and A. Herrmann Wolfgang (John Wiley & Sons, Ltd., 2017) Chap. 2, pp. 23–90 (cit. on p. 5).
[41] G. Lopez, M. Artetxe, M. Amutio, J. Alvarez, J. Bilbao, and M. Olazar, Renew. Sustain. Energy Rev. **82**, 576 (2018) (cit. on p. 5).
[42] J. Artz, T. E. Müller, K. Thenert, J. Kleinekorte, R. Meys, A. Sternberg, A. Bardow, and W. Leitner, Chem. Rev. **118**, 434 (2018) (cit. on p. 5).
[43] G. A. Olah, A. Goeppert, and G. K. S. Prakash, *Beyond Oil and Gas: The Methanol Economy*, Third ed. (Wiley-VCH, 2018) (cit. on p. 5).
[44] J. Ren, J.-P. Cao, X.-Y. Zhao, F.-L. Yang, and X.-Y. Wei, Renew. Sustain. Energy Rev. **116**, 109426 (2019) (cit. on p. 5).
[45] R. Dittmeyer, M. Klumpp, P. Kant, and G. Ozin, Nat. Commun. **10**, 1818 (2019) (cit. on p. 5).
[46] J. B. Zimmerman, P. T. Anastas, H. C. Erythropel, and W. Leitner, Science **367**, 397 (2020) (cit. on p. 5).
[47] A. J. Medford, A. C. Lausche, F. Abild-Pedersen, B. Temel, N. C. Schjødt, J. K. Nørskov, and F. Studt, Top. Catal. **57**, 135 (2014) (cit. on p. 6).
[48] K. Reuter, C. P. Plaisance, H. Oberhofer, and M. Andersen, J. Chem. Phys. **146**, 040901 (2017) (cit. on pp. 6, 20).
[49] M. Andersen, A. J. Medford, J. K. Nørskov, and K. Reuter, Angew. Chem. Int. Ed. **55**, 5210 (2016) (cit. on pp. 6, 16).
[50] M. Andersen, A. J. Medford, J. K. Nørskov, and K. Reuter, ACS Catal. **7**, 3960 (2017) (cit. on pp. 6, 16).
[51] M. Schumann, M. R. Nielsen, T. E. L. Smitshuysen, T. W. Hansen, C. D. Damsgaard, A.-C. A. Yang, M. Cargnello, J.-D. Grunwaldt, A. D. Jensen, and J. M. Christensen, ACS Catal. **11**, 5189 (2021) (cit. on pp. 6, 36).
[52] J. P. Perdew and K. Schmidt, AIP Conf. Proc. **577**, 1 (2001) (cit. on p. 6).
[53] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Second ed. (Wiley-VCH, 2001) (cit. on p. 6).
[54] T. L. Beck, Rev. Mod. Phys. **72**, 1041 (2000) (cit. on p. 6).
[55] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys.: Condens. Matter **21**, 395502 (2009) (cit. on p. 6).
[56] J. Wellendorff, K. T. Lundgaard, A. Møgelhøj, V. Petzold, D. D. Landis, J. K. Nørskov, T. Bligaard, and K. W. Jacobsen, Phys. Rev. B **85**, 235149 (2012) (cit. on p. 6).
[57] P. J. Feibelman, B. Hammer, J. K. Nørskov, F. Wagner, M. Scheffler, R. Stumpf, R. Watwe, and J. Dumesic, J. Phys. Chem. B **105**, 4018 (2001) (cit. on pp. 7, 40).
[58] F. Abild-Pedersen and M. P. Andersson, Surf. Sci. **601**, 1747 (2007) (cit. on pp. 7, 40).
[59] A. Patra, H. Peng, J. Sun, and J. P. Perdew, Phys. Rev. B **100**, 035442 (2019) (cit. on pp. 7, 40).

[60] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, Second ed. (John Wiley & Sons Ltd., 2004) (cit. on pp. 7, 9).
[61] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, J. Phys. Condens. Matter **29**, 273002 (2017) (cit. on pp. 7, 9, 10).
[62] C. Oses, E. Gossett, D. Hicks, F. Rose, M. J. Mehl, E. Perim, I. Takeuchi, S. Sanvito, M. Scheffler, Y. Lederer, O. Levy, C. Toher, and S. Curtarolo, J. Chem. Inf. Model. **58**, 2477 (2018) (cit. on pp. 9, 10, 16–18).
[63] G. Henkelman and H. Jónsson, J. Chem. Phys. **111**, 7010 (1999) (cit. on p. 10).
[64] G. Mills and H. Jónsson, Phys. Rev. Lett. **72**, 1124 (1994) (cit. on p. 10).
[65] H. Jónsson, G. Mills, and K. W. Jacobsen, "Nudged elastic band method for finding minimum energy paths of transitions," in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, C. Giovanni, and D. F. Coker (World Scientific, 1998) Chap. 16, pp. 385–404 (cit. on p. 10).
[66] G. Henkelman, B. P. Uberuaga, and H. Jónsson, J. Chem. Phys. **113**, 9901 (2000) (cit. on p. 10).
[67] S. J. Kurdziel, J. L. Lansford, and D. G. Vlachos, J. Phys. Chem. C **125**, 19780 (2021) (cit. on pp. 10, 11).
[68] F. Abild-Pedersen, J. Greeley, F. Studt, J. Rossmeisl, T. R. Munter, P. G. Moses, E. Skúlason, T. Bligaard, and J. K. Nørskov, Phys. Rev. Lett. **99**, 016105 (2007) (cit. on pp. 10, 11).
[69] B. Hammer and J. K. Nørskov, Surf. Sci. **343**, 211 (1995) (cit. on p. 10).
[70] S. Wang, V. Petzold, V. Tripkovic, J. Kleis, J. G. Howalt, E. Skúlason, E. M. Fernández, B. Hvolbæk, G. Jones, A. Toftelund, H. Falsig, M. Björketun, F. Studt, F. Abild-Pedersen, J. Rossmeisl, J. K. Nørskov, and T. Bligaard, Phys. Chem. Chem. Phys. **13**, 20760 (2011) (cit. on p. 11).
[71] A. C. Lausche, A. J. Medford, T. S. Khan, Y. Xu, T. Bligaard, F. Abild-Pedersen, J. K. Nørskov, and F. Studt, J. Catal. **307**, 275 (2013) (cit. on pp. 11, 35).
[72] J. N. Brønsted, Chem. Rev. **5**, 231 (1928) (cit. on pp. 11, 17).
[73] M. G. Evans and M. Polanyi, Trans. Faraday Soc. **34**, 11 (1938) (cit. on pp. 11, 17).
[74] A. Michaelides, Z.-P. Liu, C. J. Zhang, A. Alavi, D. A. King, and P. Hu, J. Am. Chem. Soc. **125**, 3704 (2003) (cit. on p. 11).
[75] J. E. Sutton and D. G. Vlachos, ACS Catal. **2**, 1624 (2012) (cit. on pp. 11, 17).
[76] S. Wang, V. Vorotnikov, J. E. Sutton, and D. G. Vlachos, ACS Catal. **4**, 604 (2014) (cit. on p. 11).
[77] F. Hess, J. Comput. Chem. **40**, 2664 (2019) (cit. on pp. 11, 12).
[78] M. Jørgensen and H. Grönbeck, ACS Catal. **7**, 5054 (2017) (cit. on p. 11).
[79] S. Piccinin and M. Stamatakis, Top. Catal. **60**, 141 (2017) (cit. on p. 11).

[80] E. Vignola, S. N. Steinmann, B. D. Vandegehuchte, D. Curulla, M. Stamatakis, and P. Sautet, J. Chem. Phys. **147**, 054106 (2017) (cit. on p. 11).
[81] M. Huš and A. Hellman, ACS Catal. **9**, 1183 (2019) (cit. on p. 11).
[82] K. Reuter, "First-Principles Kinetic Monte Carlo Simulations for Heterogeneous Catalysis: Concepts, Status, and Frontiers," in *Modeling and Simulation of Heterogeneous Catalytic Reactions*, edited by O. Deutschmann (John Wiley & Sons, Ltd., 2011) Chap. 3, pp. 71–111 (cit. on pp. 11, 21, 26, 27, 39).
[83] D.-J. Liu, F. Zahariev, M. S. Gordon, and J. W. Evans, J. Phys. Chem. C **120**, 28639 (2016) (cit. on p. 11).
[84] M. Pineda and M. Stamatakis, J. Chem. Phys. **147**, 024105 (2017) (cit. on p. 11).
[85] J. M. Sanchez, F. Ducastelle, and D. Gratias, Physica A **128**, 334 (1984) (cit. on p. 11).
[86] C. Stampfl, H. J. Kreuzer, S. H. Payne, H. Pfnür, and M. Scheffler, Phys. Rev. Lett. **83**, 2993 (1999) (cit. on p. 11).
[87] S. Müller, J. Phys. Condens. Matter **15**, R1429 (2003) (cit. on p. 11).
[88] Y. Zhang, V. Blum, and K. Reuter, Phys. Rev. B **75**, 235406 (2007) (cit. on pp. 11, 12).
[89] M. Stamatakis and D. G. Vlachos, J. Chem. Phys. **134**, 214115 (2011) (cit. on pp. 12, 22, 24, 40).
[90] L. Yang, A. Karim, and J. T. Muckerman, J. Phys. Chem. C **117**, 3414 (2013) (cit. on p. 12).
[91] J. Nielsen, M. d'Avezac, J. Hetherington, and M. Stamatakis, J. Chem. Phys. **139**, 224706 (2013) (cit. on pp. 12, 27).
[92] T. Hey, S. Tansley, K. Tolle, and J. Gray, *The Fourth Paradigm: Data-Intensive Scientific Discovery* (Microsoft Research, 2009) (cit. on p. 13).
[93] K. Rajan, Annu. Rev. Mater. Res. **45**, 153 (2015) (cit. on p. 13).
[94] A. Agrawal and A. Choudhary, APL Mater. **4**, 053208 (2016) (cit. on pp. 13, 14, 40).
[95] A. A. White, MRS Bull. **38**, 594 (2013) (cit. on p. 13).
[96] C. H. Ward, J. A. Warren, and R. J. Hanisch, Integr. Mater. Manuf. Innov. **3**, 292 (2014) (cit. on p. 13).
[97] S. R. Kalidindi and M. De Graef, Annu. Rev. Mater. Res. **45**, 171 (2015) (cit. on p. 13).
[98] C. Draxl and M. Scheffler, J. Phys. Mater. **2**, 036001 (2019) (cit. on pp. 13, 16).
[99] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, Fourth ed. (Pearson Education, 2021) (cit. on p. 13).
[100] P. Xu, J. Phys. Conf. Ser. **1187**, 052103 (2019) (cit. on pp. 13, 14).
[101] M. Alloghani, D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf, "A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science," in *Supervised and Unsupervised Learning for Data Science*, edited by M. W. Berry, A. Mohamed, and B. W. Yap (Springer International Publishing, 2020), pp. 3–21 (cit. on pp. 13, 14).
[102] I. H. Sarker, SN Comput. Sci. **2**, 160 (2021) (cit. on pp. 13, 14).
[103] R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, and L. M. Ghiringhelli, Phys. Rev. Materials **2**, 083802 (2018) (cit. on pp. 14, 15, 35).

[104] R. Ouyang, E. Ahmetcik, C. Carbogno, M. Scheffler, and L. M. Ghiringhelli, J. Phys. Mater. **2**, 024002 (2019) (cit. on pp. 14, 15, 35).
[105] A. Jain, G. Hautier, C. J. Moore, S. Ping Ong, C. C. Fischer, T. Mueller, K. A. Persson, and G. Ceder, Comput. Mater. Sci. **50**, 2295 (2011) (cit. on p. 16).
[106] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, APL Mater. **1**, 011002 (2013) (cit. on p. 16).
[107] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, JOM **65**, 1501 (2013) (cit. on p. 16).
[108] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, and C. Wolverton, npj Comput. Mater. **1**, 15010 (2015) (cit. on p. 16).
[109] *Citrine Informatics*, Accessed: 2022-07-31 (cit. on p. 16).
[110] D. D. Landis, J. S. Hummelshøj, S. Nestorov, J. Greeley, M. Dułak, T. Bligaard, J. K. Nørskov, and K. W. Jacobsen, Comput. Sci. Eng. **14**, 51 (2012) (cit. on p. 16).
[111] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L. W. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, and O. Levy, Comput. Mater. Sci. **58**, 227 (2012) (cit. on p. 16).
[112] L. C. Grabow, ChemCatChem **4**, 1887 (2012) (cit. on p. 16).
[113] *Automatic FLOW for Materials Discovery*, Accessed: 2022-07-31 (cit. on p. 16).
[114] S. Curtarolo, W. Setyawan, G. L. W. Hart, M. Jahnatek, R. V. Chepulskii, R. H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M. J. Mehl, H. T. Stokes, D. O. Demchenko, and D. Morgan, Comput. Mater. Sci. **58**, 218 (2012) (cit. on p. 16).
[115] F. Rose, C. Toher, E. Gossett, C. Oses, M. B. Nardelli, M. Fornari, and S. Curtarolo, Comput. Mater. Sci. **137**, 362 (2017) (cit. on p. 16).
[116] N. N. Greenwood and A. Earnshaw, *Chemistry of the Elements*, Second ed. (Butterworth-Heinemann, 1997) (cit. on p. 17).
[117] A. Bruix, J. T. Margraf, M. Andersen, and K. Reuter, Nat. Catal. **2**, 659 (2019) (cit. on pp. 18–21, 24).
[118] S. Matera, H. Meskine, and K. Reuter, J. Chem. Phys. **134**, 064713 (2011) (cit. on pp. 19, 21, 25, 27, 30).
[119] Z. W. Ulissi, A. J. Medford, T. Bligaard, and J. K. Nørskov, Nat. Commun. **8**, 14621 (2017) (cit. on p. 20).
[120] C. F. Goldsmith and R. H. West, J. Phys. Chem. C **121**, 9970 (2017) (cit. on p. 20).
[121] Q. Li, R. García-Muelas, and N. López, Nat. Commun. **9**, 526 (2018) (cit. on p. 20).
[122] J. T. Margraf and K. Reuter, ACS Omega **4**, 3370 (2019) (cit. on p. 20).
[123] S. Stocker, G. Csányi, K. Reuter, and J. T. Margraf, Nat. Commun. **11**, 5505 (2020) (cit. on p. 20).
[124] G. Henkelman and H. Jónsson, J. Chem. Phys. **115**, 9657 (2001) (cit. on p. 20).
[125] L. Xu, D. Mei, and G. Henkelman, J. Chem. Phys. **131**, 244520 (2009) (cit. on p. 20).
[126] S. T. Chill and G. Henkelman, J. Chem. Phys. **140**, 214110 (2014) (cit. on p. 20).
[127] E. W. Hansen and M. Neurock, Chem. Eng. Sci. **54**, 3411 (1999) (cit. on p. 20).
[128] E. W. Hansen and M. Neurock, Surf. Sci. **464**, 91 (2000) (cit. on p. 20).
[129] E. W. Hansen and M. Neurock, J. Catal. **196**, 241 (2000) (cit. on p. 20).

[130] K. Reuter, D. Frenkel, and M. Scheffler, Phys. Rev. Lett. **93**, 116105 (2004) (cit. on p. 20).
[131] D. Mei, P. A. Sheth, M. Neurock, and C. M. Smith, J. Catal. **242**, 1 (2006) (cit. on p. 20).
[132] K. Reuter and M. Scheffler, Phys. Rev. B **73**, 045433 (2006) (cit. on pp. 20, 22, 24).
[133] H. Meskine, S. Matera, M. Scheffler, K. Reuter, and H. Metiu, Surf. Sci. **603**, 1724 (2009) (cit. on pp. 20, 30).
[134] D. Mei, J. Du, and M. Neurock, Ind. Eng. Chem. Res. **49**, 10364 (2010) (cit. on p. 20).
[135] M. J. Hoffmann and K. Reuter, Top. Catal. **57**, 159 (2014) (cit. on pp. 20, 27).
[136] M. J. Hoffmann, M. Scheffler, and K. Reuter, ACS Catal. **5**, 1199 (2015) (cit. on pp. 20, 27).
[137] M. Andersen, C. P. Plaisance, and K. Reuter, J. Chem. Phys. **147**, 152705 (2017) (cit. on pp. 20, 22, 24, 29, 31, 35).
[138] V. J. Bukas, "Dissociation and dissipation dynamics of adsorbates at solid surfaces," PhD thesis (Technical University of Munich, 2016) (cit. on p. 21).
[139] P. Gelß, S. Matera, and C. Schütte, J. Comput. Phys. **314**, 489 (2016) (cit. on pp. 21, 25).
[140] A. F. Voter, in Radiation effects in solids, edited by K. E. Sickafus, E. A. Kotomin, and B. P. Uberuaga (2007), pp. 1–23 (cit. on p. 21).
[141] M. Karikorpi, S. Holloway, N. Henriksen, and J. K. Nørskov, Surf. Sci. Lett. **179**, L41 (1987) (cit. on pp. 22, 24).
[142] K. Reuter and M. Scheffler, Phys. Rev. B **65**, 035406 (2001) (cit. on p. 24).
[143] A. J. Medford, C. Shi, M. J. Hoffmann, A. C. Lausche, S. R. Fitzgibbon, T. Bligaard, and J. K. Nørskov, Catal. Lett. **145**, 794 (2015) (cit. on pp. 25, 37).
[144] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975) (cit. on p. 25).
[145] D. T. Gillespie, J. Comput. Phys. **22**, 403 (1976) (cit. on p. 25).
[146] A. P. J. Jansen, Comput. Phys. Commun. **86**, 1 (1995) (cit. on p. 25).
[147] M. J. Hoffmann, S. Matera, and K. Reuter, Comput. Phys. Commun. **185**, 2138 (2014) (cit. on pp. 26, 27, 37, 40).
[148] S. Piccinin and M. Stamatakis, ACS Catal. **4**, 2143 (2014) (cit. on p. 27).
[149] J. M. Lorenzi, S. Matera, and K. Reuter, ACS Catal. **6**, 5191 (2016) (cit. on p. 27).
[150] M. Seibt, "Temporal Acceleration of Kinetic Monte Carlo Simulations Applied to Heterogeneous Catalysis," Master's thesis (Technical University of Munich, 2018) (cit. on pp. 27, 37).
[151] Z. Chen, H. Wang, N. Q. Su, S. Duan, T. Shen, and X. Xu, ACS Catal. **8**, 5816 (2018) (cit. on p. 27).
[152] C. T. Campbell, Top. Catal. **1**, 353 (1994) (cit. on p. 30).
[153] C. T. Campbell, ACS Catal. **7**, 2770 (2017) (cit. on p. 30).
[154] C. Stegelmann, A. Andreasen, and C. T. Campbell, J. Am. Chem. Soc. **131**, 8077 (2009) (cit. on p. 30).
[155] A. J. Medford, *CatMAP Documentation (Release 0.2.79)*, Accessed: 2022-07-31, 2014 (cit. on p. 30).

[156] K. Reuter, Catal. Lett. **146**, 541 (2016) (cit. on p. 30).
[157] J. E. Sutton, W. Guo, M. A. Katsoulakis, and D. G. Vlachos, Nat. Chem. **8**, 331 (2016) (cit. on p. 30).
[158] B. R. Goldsmith, J. Esterhuizen, J.-X. Liu, C. J. Bartel, and C. Sutton, AIChE J. **64**, 2311 (2018) (cit. on p. 40).
[159] T. Toyao, Z. Maeno, S. Takakusagi, T. Kamachi, I. Takigawa, and K.-i. Shimizu, ACS Catal. **10**, 2260 (2020) (cit. on p. 40).

Appendices

Publication # 1

Active Site Representation in First-Principles Microkinetic Models: Data-Enhanced Computational Screening for Improved Methanation Catalysts

Martin Deimel, Karsten Reuter, and Mie Andersen
ACS Catal. **10**, 13729 (2020).
DOI: 10.1021/acscatal.0c04045

ACS Catalysis

pubs.acs.org/acscatalysis

Research Article

Active Site Representation in First-Principles Microkinetic Models: Data-Enhanced Computational Screening for Improved Methanation Catalysts

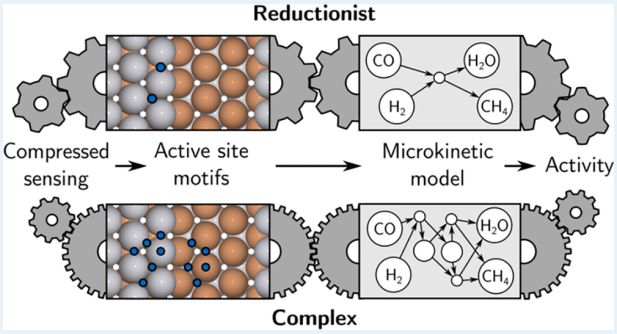
Martin Deimel, Karsten Reuter, and Mie Andersen*

Cite This: *ACS Catal.* 2020, 10, 13729–13736

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Computational screening based on first-principles microkinetic modeling has evolved into a widespread tool for catalyst discovery. Efficiently exploiting various scaling relations, this approach draws its predictive character from reliable adsorption energies, typically calculated with density-functional theory (DFT). In prevalent screening approaches, the concomitant computational costs are kept tractable through the use of reductionist microkinetic models that only resolve a minimalistic amount of active site motifs at the catalyst surface. Here, we scrutinize this common practice by systematically comparing the screening predictions for the CO methanation reaction when using microkinetic models that resolve an increasing amount of sites, up to the full consideration of all high-symmetry sites at stepped transition metal (TM) and binary TM alloy catalysts. Apart from generally overestimating the catalytic activity, the simplified models fail to identify a most promising class of layered bimetallic alloys as their insufficient representation of the catalyst surface does not allow them to correctly capture the rate-determining step. Only the full microkinetic model provides this proper mechanistic basis for the screening. The excessive amount of predictive-quality adsorption energetics required for this model is obtained from a compressed sensing descriptor that once trained readily provides these data for a new material from a single DFT calculation of the clean surface. With the current methodological advances in areas such as compressed sensing and machine learning, and the concurrent availability of cheap adsorption energetics for a wide range of possible catalyst materials, there is thus no reason to continue to use simplistic microkinetic models in computational catalyst screening.



KEYWORDS: *computational screening, heterogeneous catalysis, active site representation, density functional theory, microkinetic modeling, compressed sensing, machine learning*

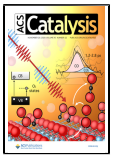
INTRODUCTION

The development of improved heterogeneous catalysts from abundant and unproblematic materials plays a crucial role for the efficient and environmentally friendly conversion of feedstock to basic and fine chemicals. Traditionally, this task comprises a vast amount of laborious laboratory tests.[1,2] With the ascent of predictive-quality theoretical methods, computational screening is increasingly used to aid in this procedure by narrowing down possible candidates for subsequent detailed testing.[3,4] In a prevalent realization, the key quantities for such computational catalyst screening are the adsorption energetics of the involved species and reaction intermediates at specific active sites on the surface. These, together with Brønsted–Evans–Polanyi relationships[4−8] for the activation energies, are the basis for the construction of microkinetic models that evaluate the kinetics of the process and predict the intrinsic catalytic activity for a candidate material.[9−12]

So far, the active site representation in such models has been limited. On the one hand, this practice arises from the desire to arrive at reductionist models that maximally condense the relevant physics. On the other hand, one also has to acknowledge that for complex reactions a wealth of new adsorption energies is needed for each additional site that is taken into account.[13,14] Since computationally demanding first-principles calculations, typically within DFT, are generally used to provide these adsorption energies at the required predictive quality, a consideration of more active sites can quickly become intractable.

With the rapid developments in advanced data science and machine learning (ML) methods, the latter limitation no

Received: September 15, 2020
Revised: October 19, 2020
Published: November 10, 2020



https://dx.doi.org/10.1021/acscatal.0c04045
ACS Catal. 2020, 10, 13729–13736

longer holds true as these techniques can substitute the DFT calculations and provide computationally undemanding and accurate energetics.[12,15−18] As an example, we here revisit the industrially important methanation reaction[19−23] and use the established Sure Independence Screening and Sparsifying Operator (SISSO) approach to identify a descriptor that allows the efficient generation of reliable adsorption energetics data for all methanation reaction intermediates at all possible adsorption sites at a wide range of TM alloy surfaces.[17,24,25] We focus here on alloys since the wealth of possible active sites offered by their surfaces provides a known means to tune the adsorption energies in order to maximize the catalytic activity.[26−29] Simultaneously, precisely this wealth of sites and the importance to properly represent it constitutes a key example of the challenges for conventional computational screening with simplified microkinetic models.

The SISSO approach instead allows us to employ a much more detailed active site representation at no additional cost. In this work, we use this to perform a theory−theory comparison on how such fine-grained microkinetic models compare with more reductionist models typically employed in the literature—not only in terms of providing the correct mechanistic understanding but also specifically with respect to the designated screening task to identify promising materials. We note that the active site representation is a choice to make in both mean-field and kinetic Monte Carlo (KMC) microkinetic models.[30,31] Our work is therefore of fundamental importance for both strands of microkinetic modeling, even if we here focus only on mean-field models. The latter are more convenient for screening purposes, as the computational cost to solve the coupled differential equations is completely negligible compared to KMC simulations or DFT calculations.

We find that models with a simpler active site representation have a general tendency to overestimate the catalytic activity. At the same time, they grossly underestimate the activity of interesting active site motifs made up of layered alloys. Alloys exhibiting these motifs would thereby be completely missed in the standard screening approach. We trace this deficiency back to the problematic approximation of reductionist microkinetic models to only consider adsorption at the most stable site of each species. In the present methanation example, this provides the wrong energetics for the rate-determining step (RDS). By also including metastable adsorption sites, a more faithful representation of the actual geometry of the RDS can be achieved.

METHODS

DFT Calculations. The DFT data set used for the SISSO descriptor identification is taken from ref 17 and enlarged to consider also the adsorbates CH_2 and CH_3 and the TMs Fe and Co, which are of relevance for the methanation reaction. The combined data set then includes adsorption energies of H, C, CH, CH_2, CH_3, CO, O, and OH on the fcc(211) facets of Co, Ni, Cu, Ru, Rh, Pd, Ag, Ir, Pt, and Au as well as of the four single-atom (SA) alloys Ag@Cu, Pt@Rh, Pd@Ir, and Au@Ni and the four AB bimetallic alloys AgPd, IrRu, PtRh, and AgAu. Figure 1 shows all considered high-symmetry sites at the stepped fcc(211) facet. Additionally, the respective adsorption energies are calculated on the bcc(210) facet of Fe and the stepped hcp(0001) facet of Co, which have similar adsorption sites to those of the fcc(211) facet (see Supporting Information Figure S1).

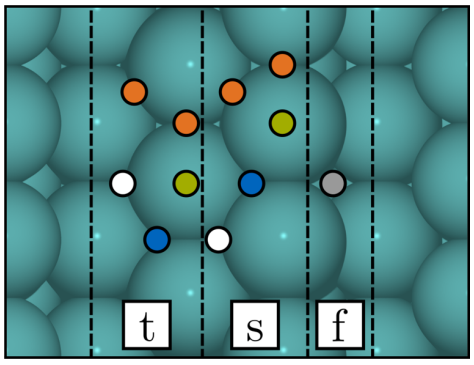


Figure 1. Top view of the stepped fcc(211) facet showing all considered high-symmetry adsorption sites. The dashed lines divide the surface area into the three coarse-grained groups: terrace (t), upper step (s) edge, and lower 4-fold (f) step edge. The colored circles indicate bridge (orange), 4-fold (gray), top (green), hcp (blue), and fcc (white) sites.

As in ref 17, the adsorption energies were calculated using the plane-wave DFT code Quantum ESPRESSO[32] with the Bayesian error estimation functional with van der Waals correlation (BEEF-vdW).[33] All fcc surface slabs are modeled using a (1 × 3) supercell with a thickness of 15 metal layers except for the alloys, which use a (1 × 1) supercell. The stepped Co hcp(0001) facet is modeled in a five-layer (1 × 3) supercell and the Fe bcc(210) facet in a 15-layer (1 × 3) supercell. The alloys generated for screening and associated primary feature determination (see below) use a (1 × 1) supercell. All calculations involving Fe, Co, and Ni include spin polarization. The plane-wave (charge density) cutoff is 500 eV (5,000 eV) except for calculations with Fe for which the cutoff is 1,000 eV (10,000 eV). The Brillouin zone of the (1 × 3) supercell was sampled with a (4 × 4) k-point grid, and equivalent k-point densities were employed for the other supercells. Further details are given in Supporting Information section S1.

Sure Independence Screening and Sparsifying Operator (SISSO). Within signal processing, compressed sensing is a tool to recover a signal from only a limited number of measurements. This methodology relies on the identification of the essential characteristics of the signal—the features.[34] SISSO[24] is such a compressed sensing method. When this is applied to a materials science problem, one is searching for the specific features that describe a physical quantity like the adsorption energy. The basis of this approach is formed by preselected primary features, e.g., coordination numbers, d-band moments, or density of states at the Fermi level, which are believed to play a role in describing the quantity of interest. Descriptors are then expressed as linear combinations of features, where the features are themselves nonlinear functions of the primary features. There are two hyperparameters in the method: one is the dimension (the number of linearly combined features), and the other is the rung (the complexity of features determined by the number of iterations in which the primary features are combined by applying algebraic/functional operators). SISSO is able to tackle huge feature spaces and to identify the best sparse solution (the solution with a certain, small number of linearly combined features) by l_0 regularization carried out in a smaller feature space selected by sure independence screening.

https://dx.doi.org/10.1021/acscatal.0c04045
ACS Catal. 2020, 10, 13729−13736

As detailed in Supporting Information section S2, we here follow the approach developed in ref 17 regarding primary features, feature construction, and the use of multitask learning[25] to identify a common descriptor for all adsorbates simultaneously. The latter allows the capture of mutually relevant features and a significant increase of the size of the training database used for descriptor identification compared to single-task learning.

Microkinetic Mean-Field Simulations. The microkinetic mean-field simulations are carried out in the steady state using the CatMAP software package.[35] In accordance with previous methanation works,[30,36] the reaction conditions are set to T = 523 K and p = 1 bar with a gas composition of 1% CO, 97% H_2, 1% CH_4, and 1% H_2O, which resembles typical reactor operation conditions. The thermochemistry of the gas-phase molecules (H_2, CO, CH_4, and H_2O) is determined within the ideal gas approximation and that of the adsorbed species and TSs within the harmonic approximation using the vibrational frequencies for Rh(211) calculated in ref 30 and the ASE thermochemistry module.[37,38] TS energies are determined using the TS- and Brønsted−Evans−Polanyi-type scaling relations established in ref 30.

In creating the reaction mechanism for a microkinetic model, there are mainly two types of coarse graining to adjust the complexity. The first one is the number of possible elementary reaction steps going from the reactants to the products. In order to obtain meaningful results, this reaction network needs to contain the major components contributing to the real pathway. The second one is the number of sites to consider in this network. As a larger number of sites naturally leads to more possible reaction steps, it is closely related to the former. Often, high-symmetry sites located within certain areas of a surface are abstracted to a site group, e.g., all sites located at a step edge or at a terrace (see Figure 1). The entire group is then represented as one effective site with adsorption energetics corresponding to the most favorable adsorption energetics of each given species at any site within the group. At the microkinetic level, only this effective site contributes to the coverage of the group and blocks the remaining high-symmetry sites within the group for adsorption. The latter site blocking is thus a crude description of lateral interactions between the adsorbates.[11,39] As detailed below, we here analyze three microkinetic models that differ in the number of groups and high-symmetry sites considered for each adsorbate at the alloy surfaces to be screened. Thereby, varying degrees of constraints are imposed on the available reaction pathways.

Note that the presented turnover frequency (TOF) is normalized to per step site. This eliminates the normalization issue to account for relative concentrations of various site types and makes all three kinetic models directly comparable to each other. In the more general case, e.g., the "tiling" model[40] could allow a rigorous normalization.

Screened Alloy Surfaces. The alloy surfaces used for screening are constructed with the help of the AFLOW database.[41] Using the CHULL module, we obtain formation enthalpies of bulk bimetallic alloy structures from all possible binary combinations of the used TMs (Fe, Co, Ni, Cu, Ru, Rh, Pd, Ag, Ir, Pt, Au).[42] The resulting convex hull plot (see Supporting Information Figure S3 for an example) shows the formation enthalpy against the TM mixing ratio and allows the identification of stable and metastable phases. As detailed in Supporting Information section S3, one challenge in determining the alloys used for the screening is the variety of different space groups to which the structures belong. A careful analysis of the convex hull plots led to the choice of the space groups 221 and 225 (fcc structure) and 187 and 164 (hcp structure). Of these, only stable and metastable structures are considered to enable a meaningful screening. Images of the different slab models used can be found in Supporting Information Figure S4.

RESULTS AND DISCUSSION

The aim of this work is to study the influence of the active site representation employed in the microkinetic model on the predicted catalyst activity and the concomitant ability to properly identify promising candidate materials in a screening procedure. The objective is thereby to scrutinize prevalent reductionist approaches and establish a best practice. To enable a systematic unbiased comparison, all microkinetic models tested draw on the same, identical pool of adsorption energetics provided by SISSO, as well as the same scaling relations for the activation energies. Differences, consequently, arise solely from the active site choices within the microkinetic models. For this theory−theory comparison, the actual accuracy of the SISSO adsorption energy predictions plays in principle only a subordinate role. Nevertheless, in order to generate a most realistic test case and simultaneously arrive at direct insights into the methanation reaction, we carefully optimize the SISSO descriptor following the approach of ref 17 and using the here established enlarged DFT data set. As detailed in Supporting Information section S2, we randomly partition each task (adsorbate) of this data set into 80% training data and 20% validation data using the scikit-learn software package[43] and optimize the two hyperparameters in the SISSO method (dimension and rung). The best prediction of the validation data set is obtained using rung 3 descriptors and levels out at dimension 5. This trend is indicative of overfitting at higher dimensions. The optimal five-dimensional rung 3 descriptor has a prediction root-mean-square error of 0.183 eV and maximum absolute error of 0.565 eV, which is fully comparable to the accuracy reached in our previous work for a smaller range of TM catalysts and reaction intermediates. The final descriptor used in the screening, see Supporting Information section S2, is obtained by fixing the hyperparameters to these optimal values and retraining the SISSO descriptor based on the entire DFT data set.

With the energetics at hand, we now introduce the microkinetic models that are used to determine the catalyst activity. Here, we compare three different models taken from published work.[30,36] According to their level of detail, they are labeled in decreasing order as complex, simple, and reductionist. For all models the surface area of the catalyst is coarse grained into specific groups indicated by the vertical dashed lines in Figure 1. Each high-symmetry site is assigned to one of these three groups termed upper step (s) edge, terrace (t), or lower 4-fold (f) step edge. The models differ in the way they resolve the true high-symmetry sites of the catalyst surface and adsorption to them in these three groups. Following refs 30 and 36, the behavior of hydrogen is approximated by adsorption on a separate "hydrogen reservoir" site (h) having the adsorption energy of the terrace site group.

The complex microkinetic model is taken from Andersen et al.[30] Here, C and CH adsorb within all three groups, whereas the other adsorbates (CO, O, OH, CH_2, CH_3, H) prefer lower coordinated sites provided in the t- and s-groups. Furthermore, all high-symmetry sites highlighted in Figure 1 are explicitly

https://dx.doi.org/10.1021/acscatal.0c04045
ACS Catal. 2020, 10, 13729−13736

Table 1. Accessible High-Symmetry Sites within Each Site Group for the Various Methanation Reaction Intermediates in the Three Microkinetic Models (See Figure 1 for an Explanation of the Different Sites and Site Groups)

	complex			simple			reductionist		
adsorbate j	s	t	f	s	t	f	s	t	f
C	all but top	all but top	ff			ff			ff
CH	all but top	all but top	ff			ff			ff
CO	all	all		br, hcp			hcp		
H		all but top			top, hcp, fcc			hcp	
O	all but top	all but top		br, hcp			hcp		
OH	all	all		br, hcp			br		
CH_2	all	all			top, hcp, fcc		br		
CH_3	all	all			top, hcp, fcc		br		

tested to find the most stable adsorption site for each adsorbate within each group. As each adsorbate may adsorb in more than one site group, the respective diffusion processes are also included. This leads to a total of 35 different elementary steps (see Supporting Information section S4).

The simple microkinetic model is taken from Lausche et al.[36] In contrast to the complex model, each adsorbate can only adsorb in one site group. Additionally, the pool of high symmetry sites within each group does not include all possibilities. C and CH adsorb exclusively on the f-group site; CO, OH, and O on bridge or hcp sites of the s-group; and H, CH_2, and CH_3 on top, hcp, or fcc sites of the t-group. Consequently, the number of elementary steps is reduced to 10, and no diffusion processes are included.

The reductionist model further reduces the site representation within each group. As Ni-based catalysts are the most important methanation catalysts for industrial applications, only the one high-symmetry site of each group with the highest stability on Ni is considered in the reductionist model. These are the f-group site for C and CH; the hcp s site for CO and O; the bridge s site for OH, CH_2, and CH_3; and the hcp t site for H.

An overview of the groups and high-symmetry sites employed in each model can be found in Table 1. Note that in a DFT-based computational screening study, the different models would imply significantly different computational costs, as they require largely different numbers of first-principles energetic data. As the most stable adsorption sites for the different reaction intermediates are not known a priori for a new catalyst material to be screened, this cost is particularly low for the reductionist model. Here, the adsorption sites on Ni are simply taken for all tested materials, thereby completely circumventing any costly calculation at different active sites. The cost of solving the complex mean-field model is also significantly higher (about 30 times) than the cost of solving the simpler models because the differential equations to be solved are stiffer in this case due to the inclusion of fast diffusion processes. However, the absolute cost is extremely low (about 0.06 core hours), which is thus negligible compared to the cost of obtaining the first-principles energetic data.

We also note that alloys contain more distinct high-symmetry sites than pure metals since sites that are identical by symmetry at the pure metal surfaces may differ in the elemental composition at alloy surfaces. For the complex and simple models, the adsorption energies were predicted at all of the possible realizations of the considered sites, and the most favorable of these sites within each site group was then used in the screening. For the reductionist model, however, the adsorption energetics were predicted at only one fixed site for all surfaces.

Using the SISSO energetics, the three microkinetic models are now employed to predict the catalytic activity for the 149 TM and TM alloy surfaces contained in our screening set. Figure 2 plots the corresponding TOFs from the complex microkinetic model against those of the simpler models (blue, simple; orange, reductionist). The solid black line represents perfect correlation. The dashed blue and orange lines are linear fits that reveal a general overestimation of the TOFs in the two simpler models. The industrial Ni catalyst and selected outliers are highlighted.

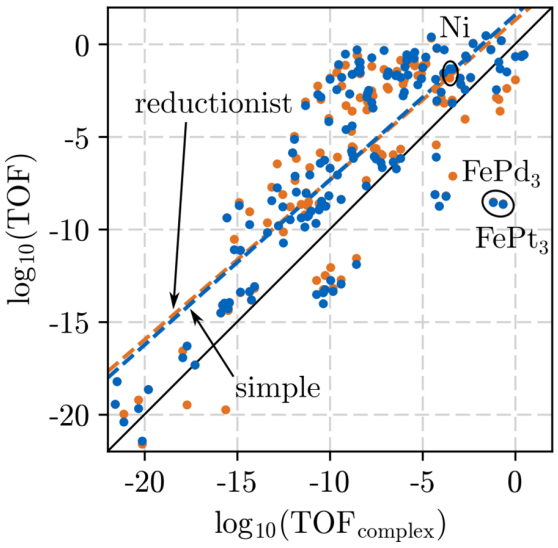


Figure 2. Correlation plot of the TOFs (product molecules per second and step site) obtained for the screened TMs and TM alloys in the complex microkinetic model against those of the simpler models (blue, simple; orange, reductionist). The solid black line represents perfect correlation. The dashed blue and orange lines are linear fits that reveal a general overestimation of the TOFs in the two simpler models. The industrial Ni catalyst and selected outliers are highlighted.

microkinetic model against the TOFs of the simple and reductionist model in blue and orange, respectively. The apparent gross correlation demonstrates that all three microkinetic models predict the same qualitative trends. Obviously, the high activity of the industrial Ni catalyst (see Supporting Information Tables S5 and S6)[20−22] is essentially identically predicted by all three models, as the site consideration of the simpler models was explicitly motivated by this material. Notwithstanding these achievements, at a closer inspection, the linear fits to the data for the simple and reductionist models reveal an overall overestimation of the catalyst activity as compared to the complex model. The correlation coefficients R^2 are about 0.6−0.7 for both regressions, and the spreads of

https://dx.doi.org/10.1021/acscatal.0c04045
ACS Catal. 2020, 10, 13729−13736

the data points around the fitted lines are more than 3 orders of magnitude. In other words, the reduced resolution of active sites in these models leads to an overall uncertainty in the predicted catalytic activity that exceeds the typical uncertainty due to the approximate DFT adsorption energetics, which is frequently discussed as a primary limitation of computational screening approaches. Even more importantly, there are also some highly active outliers like $FePt_3$ or $FePd_3$ (see their structure in Figure 3a), for which the catalytic activity is

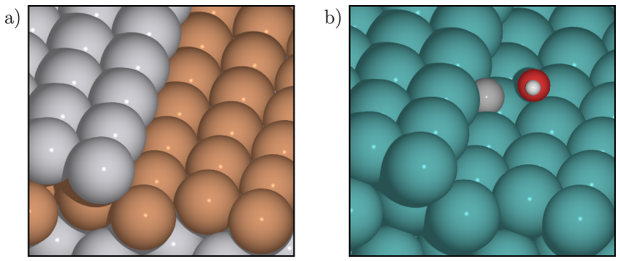


Figure 3. (a) Perspective view of the structure of the layered outliers $FePt_3$ and $FePd_3$, see Figure 2. The second layer is composed of Fe, while the top-most and lower-lying layers are composed of Pt or Pd, respectively. (b) Perspective view of the rate-determining C−OH transition state for the H-assisted CO dissociation as obtained by DFT.[30] The spatial proximity of the carbon and oxygen moieties in this late TS indicates dissociation into OH* at the t-group and C* at the f-group adsorption sites, see text.

underestimated by the simpler models by about 8 orders of magnitude and which would correspondingly be dismissed in a screening approach based on these models. The similar performance of the reductionist and simple model in these respects (see Supporting Information Figure S5) reveals that they have a similar relationship to the complex model. As they do not differ in the number of elementary steps considered, but only in the choice of the active site representation, this is a strong indication that both models lack similar information, which is included in the complex model and which is particularly pronounced for the outliers.

In order to analyze why the activities from the complex model are different from the simpler models, we determine the RDSs for each screened alloy and for each applied model by calculating Campbell's degree of rate control,[44] X_j:

$$X_j = \left(\frac{\partial \log(\mathrm{TOF})}{\partial(-G_j/k_\mathrm{B}T)} \right)_{G_{n\neq j}}$$

Here, G_j is the free energy of formation of an intermediate or transition state j, k_B is the Boltzmann constant, and T is the temperature. All other formation energies $n \neq j$ are kept constant, leading to a measure of the influence of the formation energy of a single species on the TOF. Here, we define the RDSs as those where the TS has a high DRC. The RDSs play the most important role for the TOF, and thus any difference in the energetics of the RDSs employed in the three microkinetic models (due to the varying active site representation) will directly influence the TOF. In contrast, differences in the energetics of all other reaction steps are less likely to influence the TOF. Note, though, that influences are in principle possible since microkinetic models are highly nonlinear, whereas the DRC is formulated as a linear response theory, meaning that the result is only valid locally in parameter space.

For the vast majority of the screened catalysts, we identify the H-assisted CO dissociation

$$\mathrm{CO}^* + \mathrm{H}^* \rightleftharpoons \mathrm{C{-}OH}^* + * \rightarrow \mathrm{C}^* + \mathrm{OH}^*$$

to be the dominant RDS. The full DRC analysis can be found in Supporting Information Figure S6. The geometry of the TS in this step (Figure 3b) is very similar to that of the final state; i.e., the C is adsorbed near the f site and the OH is adsorbed near the t site. However, in the simple and reductionist models, adsorption of OH at a t site is not included, since for the pure metals the s site is generally the more favorable OH adsorption site. As an example, we show in Table 2 the SISSO adsorption energies of the species involved in the RDS for Ni, which as mentioned above is one of the materials for which the activity is only slightly overestimated by the simpler models. For this material, OH adsorption at the s site is by 0.5 eV more favorable than adsorption at the t site. The overestimated TOF in the simpler models is thus caused by the limitation to consider only the more favorable energetics of OH at an s site, which in turn also influences the energetics of the TS, since it scales with the final state energetics through the employed TS scaling relation. The resulting decreased activation energy of the RDS in the simpler models can also be seen in the energy diagrams in Figure 4a. Note that also the CO adsorption sites are different in the three compared models, but this plays a smaller role since the t and s sites have generally rather similar CO adsorption energies, cf. the example of Ni in Table 2.

Having understood the general trend, we finally analyze why certain materials are prominent outliers with a significantly underestimated TOF in the simpler models. This analysis turns out to be identical for the two outliers highlighted in Figure 2, and we therefore focus here only on the $FePt_3$ alloy. Compared to the pure metals and most other alloys, the layered structure

Table 2. SISSO-Predicted Adsorption Energies (in eV) of CO*, H*, C−OH*, C*, and OH* at Ni and $FePt_3$ at the Sites Involved in the RDS in the Three Compared Models, Together with the Correspondingly Predicted TOFs, See Text[a]

Catalyst	Model	CO*		+ H* ⇌	C-OH*	→ C* +		OH*		TOF
		s	t	h	f	f	s	t		$[s^{-1}\ \mathrm{step\ site}^{-1}]$
Ni	Reductionist	0.933	-	−0.230	2.052	1.582	−0.402	-		$1.55 \cdot 10^{-2}$
	Simple	0.896	-	−0.230	2.052	1.582	−0.402	-		$4.96 \cdot 10^{-2}$
	Complex	-	1.054	−0.230	2.488	1.582	-	0.102		$2.87 \cdot 10^{-4}$
$FePt_3$	Reductionist	1.399	-	−0.158	3.265	1.879	0.704	-		$2.46 \cdot 10^{-9}$
	Simple	1.399	-	−0.188	3.265	1.879	0.704	-		$2.31 \cdot 10^{-9}$
	Complex	-	1.263	−0.188	2.376	1.879	-	−0.324		$2.16 \cdot 10^{-1}$

[a]For the Ni catalyst, the different site resolution in the three models leads only to a modest variation of the predicted activity. In contrast, for the layered $FePt_3$ catalyst, the activity is dramatically underestimated by about eight orders of magnitude by the simpler models.

https://dx.doi.org/10.1021/acscatal.0c04045
ACS Catal. 2020, 10, 13729−13736

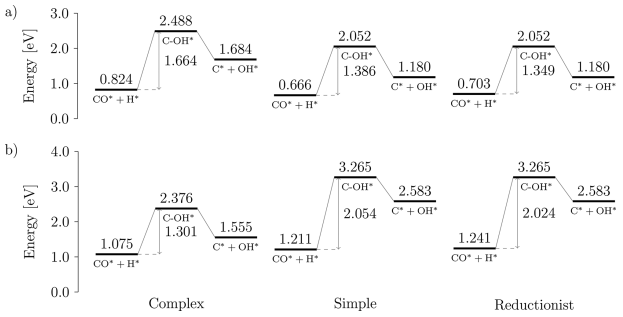


Figure 4. Energy diagrams of the RDSs in the three compared models using the SISSO adsorption energies at Ni (a) and $FePt_3$ (b). The activation energies are depicted by the vertical lines.

of the $FePt_3$ alloy (see Figure 3a) inverts the energetic ordering of the considered OH adsorption sites. Whereas for Ni, OH adsorption is more favorable at the undercoordinated s site, it is more favorable at the t site for the $FePt_3$ alloy, cf. Table 2. The reason for this is that Fe (which makes up the majority of the considered high-symmetry t sites) interacts much more strongly with OH than Pt (which makes up the considered s sites) in the $FePt_3$ alloy. This structural motif, which is unique to the layered alloys, thus allows for a particularly favorable final state (and thereby also TS) in the RDS, leading to a lowered activation energy (cf. Figure 4b) and concomitantly to an increased TOF. In general, such insight into the most stable adsorption sites for a particular catalyst surface is not available before the screening but only becomes apparent once all of the different site possibilities have been assessed, as is done here in the complex model based on the cheap energetics provided by the SISSO descriptor.

It is worth noting that the promising aspects of layered bimetallic catalysts for CO methanation have already been discussed based on DFT and microkinetic simulations.[45] However, this previous study investigated only (nine) layered bimetallic catalysts and did not consider the stability of these materials. Here, we considered a large pool of different alloy surfaces, taking care to include only materials with a stable or metastable bulk structure. There is thus no bias for a particular structural motif. In contrast, it is an intrinsic outcome of the screening that layered structures are identified as promising CO methanation catalysts. Furthermore, this result is only found when using a complex reaction model that takes into account all possible adsorption sites for the reaction intermediates.

CONCLUSIONS

We analyzed the impact of the active site representation within mean-field microkinetic models of CO methanation on elemental TMs and bimetallic TM alloys. For this, we compared the results obtained from three models of different complexity. The required energetic input was obtained from a descriptor identified using the compressed sensing method SISSO. To this end, a DFT-calculated data set of adsorption energies at elemental TM surfaces and selected alloy surfaces served as training for the descriptor identification. The descriptor is a function of properties of the clean alloy surface only and thus allows for obtaining adsorption energies of all reaction intermediates at all high-symmetry sites of the complex, stepped alloy surfaces, at the cost of a single DFT calculation of the clean surface. In the here considered example, all reaction intermediates are monodentate species. However, we see no fundamental obstacles preventing the generalization to consider also more complex adsorbates with possible bidentate adsorption motifs, as long as the calculation of the primary features entering the descriptor are appropriately defined and a sufficiently diverse training data set is employed.

The main result of our theory−theory comparison is that simpler microkinetic models previously employed in the literature generally tend to overestimate the TOF of elemental TMs and most bimetallic TM alloys due to an overly simplified representation of the adsorption sites of the reaction intermediates involved in the RDS. In contrast, the simpler models significantly underestimate the TOFs of active site motifs made up of layered bimetallic alloys, in which the energetic ordering of the considered adsorption sites for the crucial OH intermediate is reversed. These structures would thereby be completely missed if carrying out the screening using simple literature microkinetic models. The right mechanistic understanding and—consequently—accurate screening results can only be obtained using the complex model. On the basis of its detailed active site representation, we find that layered bimetallic alloys are particularly promising catalysts when the lower layer is made up of a TM that interacts strongly with OH such as Fe and the upper layer is made up of a less reactive metal such as Pd or Pt.

Overall, our results demonstrate that advanced data science and machine learning methods render the use of reductionist microkinetic models obsolete, as the cost to obtain the required adsorption energetics is no longer a bottleneck. The approach demonstrated here allows for accurate screening of complex catalyst materials at very low computational cost. We considered here the important CO methanation reaction as an example, but in general our approach is not limited to a specific reaction or materials class.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.0c04045.

Additional details on DFT, SISSO, the screened alloy surfaces, and the compared microkinetic models (PDF)

AUTHOR INFORMATION

Corresponding Author

Mie Andersen − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, 85747 Garching, Germany;* orcid.org/0000-0002-9943-1534; Email: mie.andersen@ch.tum.de

Authors

Martin Deimel − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, 85747 Garching, Germany*

Karsten Reuter − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, 85747 Garching, Germany; Fritz-Haber-Institut der Max-Planck-Gesellschaft, 14195 Berlin, Germany;* orcid.org/0000-0001-8473-8659

Complete contact information is available at: https://pubs.acs.org/10.1021/acscatal.0c04045

https://dx.doi.org/10.1021/acscatal.0c04045
ACS Catal. 2020, 10, 13729−13736

Notes

The authors declare no competing financial interest.

Input files for the CatMAP calculations for the three models together with the primary features and geometries used to obtain the predicted energetics are available at https://github.com/m-deimel/active_site_representation.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge support from the Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities (www.lrz.de) and the Jülich Supercomputing Centre (www.fz-juelich.de/ias/jsc). Partial funding came from the Solar Technologies Go Hybrid initiative of the State of Bavaria.

■ REFERENCES

(1) George, S. M. Introduction: Heterogeneous Catalysis. *Chem. Rev.* **1995**, *95*, 475−476.
(2) Ertl, G. Reactions at Surfaces: From Atoms to Complexity (Nobel Lecture). *Angew. Chem., Int. Ed.* **2008**, *47*, 3524−3535.
(3) Broadbelt, L. J.; Snurr, R. Q. Applications of molecular modeling in heterogeneous catalysis research. *Appl. Catal., A* **2000**, *200*, 23−46.
(4) Nørskov, J. K.; Studt, F.; Abild-Pedersen, F.; Bligaard, T. *Fundamental Concepts in Heterogeneous Catalysis*; John Wiley & Sons, Inc., 2014.
(5) Brønsted, J. N. Acid and Basic Catalysis. *Chem. Rev.* **1928**, *5*, 231−338.
(6) Evans, M. G.; Polanyi, M. Inertia and driving force of chemical reactions. *Trans. Faraday Soc.* **1938**, *34*, 11−24.
(7) Pallassana, V.; Neurock, M. Electronic Factors Governing Ethylene Hydrogenation and Dehydrogenation Activity of Pseudomorphic PdML/Re(0001), PdML/Ru(0001), Pd(111), and PdML/Au(111) Surfaces. *J. Catal.* **2000**, *191*, 301−317.
(8) Michaelides, A.; Liu, Z.-P.; Zhang, C. J.; Alavi, A.; King, D. A.; Hu, P. Identification of General Linear Relationships between Activation Energies and Enthalpy Changes for Dissociation Reactions at Surfaces. *J. Am. Chem. Soc.* **2003**, *125*, 3704−3705.
(9) Bligaard, T.; Nørskov, J.; Dahl, S.; Matthiesen, J.; Christensen, C.; Sehested, J. The Brønsted-Evans-Polanyi relation and the volcano curve in heterogeneous catalysis. *J. Catal.* **2004**, *224*, 206−217.
(10) Sutton, J. E.; Vlachos, D. G. A Theoretical and Computational Analysis of Linear Free Energy Relations for the Estimation of Activation Energies. *ACS Catal.* **2012**, *2*, 1624−1634.
(11) Reuter, K. Ab Initio Thermodynamics and First-Principles Microkinetics for Surface Catalysis. *Catal. Lett.* **2016**, *146*, 541−563.
(12) Bruix, A.; Margraf, J. T.; Andersen, M.; Reuter, K. First-principles-based multiscale modelling of heterogeneous catalysis. *Nature Catal.* **2019**, *2*, 659−670.
(13) Nørskov, J. K.; Abild-Pedersen, F.; Studt, F.; Bligaard, T. Density functional theory in surface chemistry and catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 937−943.
(14) Reuter, K.; Plaisance, C. P.; Oberhofer, H.; Andersen, M. Perspective: On the active site model in computational catalyst screening. *J. Chem. Phys.* **2017**, *146*, 040901.
(15) 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.
(16) Medford, A. J.; Kunz, M. R.; Ewing, S. M.; Borders, T.; Fushimi, R. Extracting Knowledge from Data through Catalysis Informatics. *ACS Catal.* **2018**, *8*, 7403−7429.
(17) Andersen, M.; Levchenko, S. V.; Scheffler, M.; Reuter, K. Beyond Scaling Relations for the Description of Catalytic Materials. *ACS Catal.* **2019**, *9*, 2752−2759.
(18) Ulissi, Z. W.; Medford, A. J.; Bligaard, T.; Nørskov, J. K. To address surface reaction network complexity using scaling relations machine learning and DFT calculations. *Nat. Commun.* **2017**, *8*, 14621.
(19) Sabatier, P.; Senderens, J.-B. Nouvelles synthèses du méthane. *Comptes Rendus Acad. Sci.* **1902**, *134*, 514−516.
(20) Fischer, F.; Tropsch, H.; Dilthey, P. Über die Reduktion von Kohlenoxyd zu Methan an verschiedenen Metallen. *Brennst.-Chemie* **1925**, *6*, 265−271.
(21) Mills, G. A.; Steffgen, F. W. Catalytic Methanation. *Catal. Rev.: Sci. Eng.* **1974**, *8*, 159−210.
(22) Vannice, M. The catalytic synthesis of hydrocarbons from H_2CO mixtures over the group VIII metals: I. The specific activities and product distributions of supported metals. *J. Catal.* **1975**, *37*, 449−461.
(23) Rönsch, S.; Schneider, J.; Matthischke, S.; Schlüter, M.; Götz, M.; Lefebvre, J.; Prabhakaran, P.; Bajohr, S. Review on methanation From fundamentals to current projects. *Fuel* **2016**, *166*, 276−296.
(24) 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*, 083802.
(25) Ouyang, R.; Ahmetcik, E.; Carbogno, C.; Scheffler, M.; Ghiringhelli, L. M. Simultaneous learning of several materials properties from incomplete databases with multi-task SISSO. *J. Phys. Mater.* **2019**, *2*, 024002.
(26) Vojvodic, A.; Nørskov, J. K. New design paradigm for heterogeneous catalysts. *Natl. Sci. Rev.* **2015**, *2*, 140−143.
(27) Mamun, O.; Winther, K. T.; Boes, J. R.; Bligaard, T. High-throughput calculations of catalytic properties of bimetallic alloy surfaces. *Sci. Data* **2019**, *6*, 76.
(28) Andersson, M. P.; Bligaard, T.; Kustov, A.; Larsen, K. E.; Greeley, J.; Johannessen, T.; Christensen, C. H.; Nørskov, J. K. Toward computational screening in heterogeneous catalysis: Pareto-optimal methanation catalysts. *J. Catal.* **2006**, *239*, 501−506.
(29) Andersen, M.; Medford, A. J.; Nørskov, J. K.; Reuter, K. Scaling-Relation-Based Analysis of Bifunctional Catalysis: The Case for Homogeneous Bimetallic Alloys. *ACS Catal.* **2017**, *7*, 3960−3967.
(30) 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.
(31) Andersen, M.; Panosetti, C.; Reuter, K. A Practical Guide to Surface Kinetic Monte Carlo Simulations. *Front. Chem.* **2019**, *7*, 202.
(32) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.
(33) 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.
(34) Ghiringhelli, L. M.; Vybiral, J.; Ahmetcik, E.; Ouyang, R.; Levchenko, S. V.; Draxl, C.; Scheffler, M. Learning physical descriptors for materials science by compressed sensing. *New J. Phys.* **2017**, *19*, 023017.
(35) 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.
(36) Lausche, A. C.; Medford, A. J.; Khan, T. S.; Xu, Y.; Bligaard, T.; Abild-Pedersen, F.; Nørskov, J. K.; Studt, F. On the effect of coverage-dependent adsorbate-adsorbate interactions for CO methanation on transition metal surfaces. *J. Catal.* **2013**, *307*, 275−282.

https://dx.doi.org/10.1021/acscatal.0c04045
ACS Catal. 2020, 10, 13729−13736

(37) Bahn, S. R.; Jacobsen, K. W. An object-oriented scripting interface to a legacy electronic structure code. *Comput. Sci. Eng.* **2002**, *4*, 56−66.
(38) Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; Hermes, E. D.; Jennings, P. C.; Jensen, P. B.; Kermode, J.; Kitchin, J. R.; Kolsbjerg, E. L.; Kubal, J.; Kaasbjerg, K.; Lysgaard, S.; Maronsson, J. B.; Maxson, T.; Olsen, T.; Pastewka, L.; Peterson, A.; Rostgaard, C.; Schiøtz, J.; Schütt, O.; Strange, M.; Thygesen, K. S.; Vegge, T.; Vilhelmsen, L.; Walter, M.; Zeng, Z.; Jacobsen, K. W. The atomic simulation environment − a Python library for working with atoms. *J. Phys.: Condens. Matter* **2017**, *29*, 273002.
(39) Zhang, Y.; Blum, V.; Reuter, K. Accuracy of first-principles lateral interactions: Oxygen at Pd(100). *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007**, *75*, 235406.
(40) Herschlag, G. J.; Mitran, S.; Lin, G. A consistent hierarchy of generalized kinetic equation approximations to the master equation applied to surface catalysis. *J. Chem. Phys.* **2015**, *142*, 234703.
(41) Curtarolo, S.; Setyawan, W.; Hart, G. L.; Jahnatek, M.; Chepulskii, R. V.; Taylor, R. H.; Wang, S.; Xue, J.; Yang, K.; Levy, O.; Mehl, M. J.; Stokes, H. T.; Demchenko, D. O.; Morgan, D. AFLOW: An automatic framework for high-throughput materials discovery. *Comput. Mater. Sci.* **2012**, *58*, 218−226.
(42) Oses, C.; Gossett, E.; Hicks, D.; Rose, F.; Mehl, M. J.; Perim, E.; Takeuchi, I.; Sanvito, S.; Scheffler, M.; Lederer, Y.; Levy, O.; Toher, C.; Curtarolo, S. AFLOW-CHULL: Cloud-Oriented Platform for Autonomous Phase Stability Analysis. *J. Chem. Inf. Model.* **2018**, *58*, 2477−2490.
(43) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.
(44) Stegelmann, C.; Andreasen, A.; Campbell, C. T. Degree of Rate Control: How Much the Energies of Intermediates and Transition States Control Rates. *J. Am. Chem. Soc.* **2009**, *131*, 8077−8082.
(45) Wang, Z.; Wang, H.-F.; Hu, P. Possibility of designing catalysts beyond the traditional volcano curve: a theoretical framework for multi-phase surfaces. *Chem. Sci.* **2015**, *6*, 5703−5711.

https://dx.doi.org/10.1021/acscatal.0c04045
ACS Catal. 2020, 10, 13729−13736

Supporting Information: Active Site Representation in First-Principles Microkinetic Models: Data-Enhanced Computational Screening for Improved Methanation Catalysts

Martin Deimel,[†] Karsten Reuter,[†,‡] and Mie Andersen[*,†]

†*Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany*

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

E-mail: mie.andersen@ch.tum.de

S1 Additional DFT computational details

In enlarging the DFT data set from ref 1, the following settings were used in full compliance with the existing calculations. We employed ultrasoft pseudopotentials and a vacuum region of 16 Å perpendicular to the surface as well as a dipole correction.[2] Pseudopotentials for Fe, Ni, Cu, Rh, Pd, Ag, and Ir were generated using the "atomic" code by A. Dal Corso (v.5.0.2 svn rev. 9415), for Co using the Vanderbilt code version 7.0.0, and for Ru using the Vanderbilt code version 7.3.5. Surface relaxations with and without an adsorbate used supercell sizes of (1×3) for the pure metal slabs and (1×1) for all alloys. Depending on the supercell size, we used $\boldsymbol{k}$-point grids of (4×4) for the pure metal fcc, SA alloy fcc, and bcc slabs, (2×6) for the SG225 fcc alloys, (4×6) for the SG221 and AB fcc alloys, (2×4) for the Co hcp slab, and (2×12) for the hcp alloy structures to ensure an equivalent $\boldsymbol{k}$-point density. For the relaxations of the fcc(211) and bcc(210) facets the bottom nine layers and for the stepped hcp(0001) the bottom three layers were fixed in their bulk truncated positions, while the layers above were relaxed until reaching a maximum force threshold of 0.03 eV/Å.

In compliance with ref 1, initial adsorption geometries of CH_2 and CH_3 were created on the basis of optimized unit cells by placing the C atom perpendicularly above the local surface plane of a relaxed slab at a distance of about 2 Å from the TM atoms comprising the site. For CH_2 (CH_3) an H was placed at a distance of about 1.1 Å from the base atom C on the prolonged perpendicular axis. This axis was tilted at the C atom about 60° and the remaining H atoms were placed equally distanced from each other by rotating the tilted vector about 180° (120°) around the perpendicular axis. In case the surface slab reconstructed during the relaxation, an adsorbate moved from the initial site, or decomposed, the calculation was discarded. A slab was considered reconstructed if any atom moved more than 1 Å from its initial position and an adsorbate was considered decomposed if the distance of the adsorbing atom to any other atom exceeded 1.7 Å. Finally, an adsorbate was considered to have moved if any distance of the coordinating adsorbate atom to the coordinating site atoms deviated by more than 10% (in the case where all metal atoms are of the same element) or 20%

(in the case where the site is comprised of more than one element), or if any distance to a neighboring metal atom became smaller than 110% of any of the adsorbate / site atom distances. All high-symmetry sites shown in Figure S1 were tested. Furthermore, a series of different orientations of the H atoms in CH_2 and CH_3 were tested for each site, and the most stable configuration was chosen for the data set. An overview of the data set is shown in Table S1. The resulting adsorption energies correspond to formation energies referenced to gaseous CH_4, H_2, and H_2O.

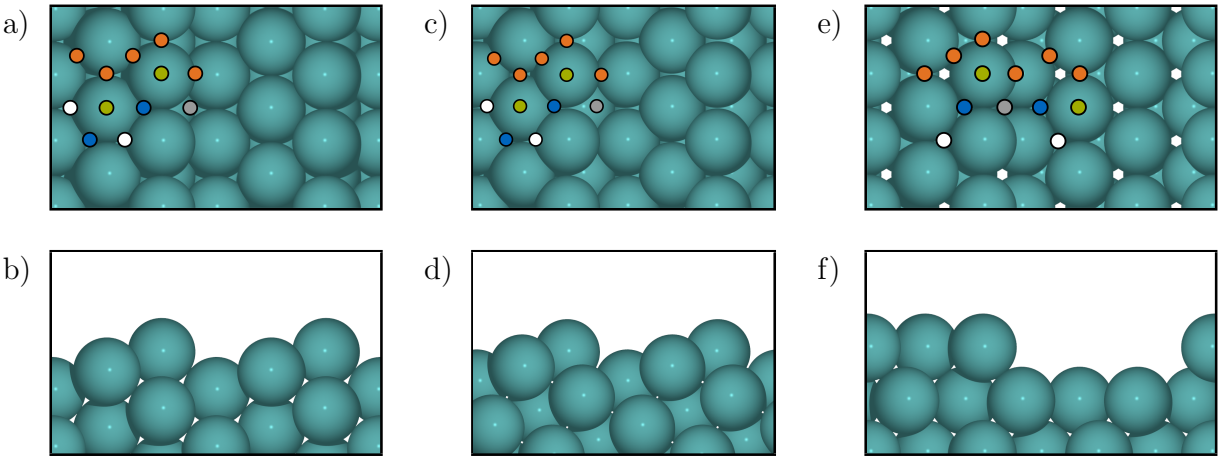


Figure S1: Depiction of the different surface facets with the corresponding tested high-symmetry sites for CH_2 and CH_3. a) and b) show the top and side view of the fcc(211) facet, c) and d) of the bcc(210) facet, and e) and f) of the stepped hcp(0001) facet. The colored circles indicate bridge (orange), 4-fold (gray), top (green), hcp (blue), and fcc (white) sites. Note that the site labels hcp and fcc of the bcc(210) facet are adapted from the fcc(211) facet.

For constructing the SISSO descriptors (see next section) we used the same list of primary features as in ref 1 (see Table S2). Details on how these primary features are calculated can be found in the original publication. For the new TMs Co and Fe we list the primary features of the classes "Atomic", "Bulk", and "Surface" in Table S3. The bulk_{nnd} of alloy structures is calculated as the average nearest neighbor distance of all atoms in the bulk unit cell and the work function is determined by the maximum average potential in the vacuum region between the periodic images. V^2_{ad}, PE, IP, EA, and r_d are calculated as the average of the atoms within the site ensemble. The calculation of the projected density of states used $\boldsymbol{k}$-

Table S1: Overview of the number of stable and metastable adsorption sites identified for the enlarged data set.

Material class	Site Class	C	O	H	CH	CO	OH	CH_2	CH_3	Total
Metals	top	2	4	3	3	24	23	23	23	105
	bridge	5	14	15	13	42	52	39	29	209
	fcc & hcp	40	47	45	40	33	34	31	16	286
	4-fold	12	3	5	12	5	-	8	1	46
	total	59	68	68	68	104	109	101	69	**646**
SA alloys	top	-	-	-	-	16	14	15	15	60
	bridge	4	7	8	8	28	43	35	19	152
	fcc & hcp	26	28	25	25	20	14	23	4	165
	4-fold	7	-	2	6	1	-	3	-	19
	total	37	35	35	39	65	71	76	38	**396**
AB alloys	top	-	-	-	-	16	10	13	16	55
	bridge	3	7	8	7	23	52	41	6	147
	fcc & hcp	20	30	17	29	6	9	8	-	119
	4-fold	8	-	2	8	2	-	4	-	24
	total	31	37	27	44	47	71	66	22	**345**

point grids of (14×14) for the pure metal fcc and bcc slabs, (7×21) for the SG225 fcc alloys, (14×21) for the SG221 fcc alloys, (7×14) for the Co hcp slab, and (7×42) for the hcp alloy structures.

Table S2: Primary features for the feature construction within SISSO taken from ref 1.

Class	Name	Abbreviation	Unit
Atomic	Pauling electronegativity	PE	–
	Ionization potential	IP	eV
	Electron affinity	EA	eV
Bulk	Nearest neighbor distance	$\mathrm{bulk_{nnd}}$	Å
	Radius of *d*-orbitals	r_d	Å
	Coupling matrix element squared	$\mathrm{V_{ad}^2}$	–
Surface	Work function	W	eV
Site	Number of atoms in ensemble	$\mathrm{site_{no}}$	#atoms
	Coordination number	CN	#atoms
	Nearest neighbor distance	$\mathrm{site_{nnd}}$	Å
	d-band center	ε_d	eV
	d-band width	W_d	eV
	d-band filling	f_d	#states
	sp-band filling	f_{sp}	#states
	Density of *d*-states at Fermi level	DOS_d	eV^{-1}
	Density of *sp*-states at Fermi level	DOS_{sp}	eV^{-1}

Table S3: Primary features of the classes "Atomic", "Bulk" and "Surface" for the added TMs Co and Fe.

	PE[a]	IP[b]	EA[b]	r_d [c]	$\mathrm{V_{ad}^2}$ [d]	$\mathrm{bulk_{nnd}}$	W
$\mathrm{Co_{hcp}}$	1.88	7.881	0.662	0.76	1.34	2.504	4.819
$\mathrm{Co_{fcc}}$	1.88	7.881	0.662	0.76	1.34	2.509	4.828
$\mathrm{Fe_{bcc}}$	1.83	7.902	0.153	0.80	1.59	2.476	4.430

[a] From ref 3 [b] From ref 4 [c] From ref 5 [d] From ref 6

S2 Additional SISSO computational details

All calculations for the descriptor identification were performed with version 3.0 of the SISSO code.[7] The primary features used are given in Table S2 and the set of algebraic/functional operations for constructing the feature spaces of rung 1, 2, and 3 is

$$\hat{\mathbf{H}}^{(m)} = \{I, +, -, \times, \div, \exp, \exp^{-1}, \log, ^{-1}, ^{2}, ^{3}, \sqrt{}, \sqrt[3]{}, \mathrm{scd}\},$$

with the standard Cauchy distribution $\mathrm{scd}(x) = \frac{1}{\pi(1+x^2)}$. The feature spaces of higher rungs are constructed consecutively from the previous rung. In order to obtain physically meaningful features, application of a binary operator from the above set is only performed on dimensionally suitable (primary) feature combinations. Sure independence screening (SIS) narrows down the created feature space to a user-defined value as the subsequent ℓ_0 regularization step is NP-hard, leading to an exponential increase in computational cost with the size of the feature (sub)space. Here we used an SIS value of 150 for the lower dimensions, which are anyway uninteresting in terms of the accuracy of the predictions of the resulting descriptors. Beyond 4 dimensions we used the highest SIS value that is still computationally tractable (SIS value of 50 for 5 dimensions and SIS value of 20 for 6 dimensions).

If the measurements can be split into sets of common origin, multitask learning enhances generalization of the resulting model by solving all tasks simultaneously. Here, a single task is the prediction of the adsorption energetics of one adsorbate. Multitask learning leads to a common model that captures mutually relevant features and increases the size of the effective training data set for each adsorbate compared to single-task learning. At a given dimension, this thereby reduces the risk of overfitting. In order to obtain a representative training and validation set of all available data, a stratified approach was applied, splitting the data set of each adsorbate randomly into 80% training and 20% validation data using the function `model_selection.StratifiedKFold` of the scikit-learn software package.[8]

The resulting root-mean-square error (RMSE) for each rung and dimension is plotted in

Figure S2. The best prediction of the validation data set is obtained using rung 3 descriptors and levels out at dimension 5. This trend is indicative of overfitting at higher dimensions. The optimal five-dimensional rung 3 descriptor has a prediction root-mean-square error of 0.183 eV and maximum absolute error of 0.565 eV. Having established the optimal hyperparameters, we used these to retrain a new descriptor based on the entire data set. The final descriptor is of the form

$$
\begin{aligned}
E_{\mathrm{ads}}^{j} = c_1^j \cdot & \left(\frac{\mathrm{V}_{\mathrm{ad}}^2 \cdot (\mathrm{W}_d + \varepsilon_d)}{\mathrm{W}} - \frac{\mathrm{scd}(\mathrm{DOS}_d) \cdot \mathrm{CN}}{\mathrm{site}_{\mathrm{no}}} \right) \\
& + c_2^j \cdot \left(\frac{\mathrm{V}_{\mathrm{ad}}^2 \cdot \mathrm{f}_{sp} \cdot \mathrm{bulk}_{\mathrm{nnd}}}{\mathrm{r}_d} - \mathrm{PE} \cdot \mathrm{f}_d \cdot \log(\mathrm{DOS}_{sp}) \right) \\
& + c_3^j \cdot \frac{(\mathrm{EA} + 2 \cdot \mathrm{W}_d - \mathrm{IP}) \cdot \mathrm{W}_d \cdot \log(\mathrm{site}_{\mathrm{no}})}{\mathrm{IP}} \\
& + c_4^j \cdot \frac{(\mathrm{scd}(\mathrm{site}_{\mathrm{no}}) - \mathrm{scd}(\mathrm{r}_d)) \cdot \varepsilon_d}{\mathrm{f}_d^3 \cdot \mathrm{site}_{\mathrm{no}}} \\
& + c_5^j \cdot \frac{\mathrm{IP} \cdot \mathrm{site}_{\mathrm{no}}}{\exp(\mathrm{site}_{\mathrm{no}}) \cdot \left(\frac{\varepsilon_d}{\mathrm{EA}} - \frac{\mathrm{CN}}{\mathrm{site}_{\mathrm{no}}} \right)} \\
& + c_0^j,
\end{aligned}
$$

where j denotes the adsorbate. The fitting coefficients $c_0^j \cdots c_5^j$ are specific to each adsorbate (see Table S4) with units ensuring that the result (the predicted adsorption energy) is in eV.

Table S4: Fitting coefficients for each of the adsorbates (tasks) of the final SISSO descriptor trained on the entire DFT data set. The units of the fitting coefficients c_j depend on the respective feature units and ensures that the unit of the predicted adsorption energy is eV. The root-mean-square error (RMSE) and maximum absolute error (MaxAE) given for each adsorbate are the training errors of the final descriptor. All values are rounded to three decimal digits of a common order of magnitude.

Adsorbate j	c_1 [eV]	c_2 $[\frac{\mathrm{eV}}{\#\mathrm{states}}]$	c_3 [-]	c_4 $[\#\mathrm{states}^3 \cdot \#\mathrm{atoms}]$	c_5 $[\frac{1}{\#\mathrm{atoms}}]$	c_0 [eV]	RMSE [eV]	MaxAE [eV]
C	-1.264	$-0.178 \cdot 10^{-2}$	1.486	$8.005 \cdot 10^{3}$	1.154	2.522	0.211	0.547
CH	-1.102	$-0.970 \cdot 10^{-2}$	1.223	$0.645 \cdot 10^{3}$	1.264	2.857	0.173	0.601
CO	-0.446	$0.813 \cdot 10^{-2}$	-0.054	$2.970 \cdot 10^{3}$	0.657	0.500	0.182	0.602
H	-0.234	$0.670 \cdot 10^{-2}$	0.385	$1.018 \cdot 10^{3}$	1.130	-0.036	0.092	0.214
O	-0.744	$5.699 \cdot 10^{-2}$	0.618	$-0.801 \cdot 10^{3}$	0.413	-0.996	0.159	0.506
OH	-0.231	$4.969 \cdot 10^{-2}$	0.168	$0.839 \cdot 10^{3}$	1.316	-1.364	0.149	0.413
CH_2	-0.515	$0.868 \cdot 10^{-2}$	0.545	$1.187 \cdot 10^{3}$	3.200	2.116	0.177	0.503
CH_3	-0.173	$2.083 \cdot 10^{-2}$	0.042	$3.130 \cdot 10^{3}$	1.851	0.206	0.161	0.419

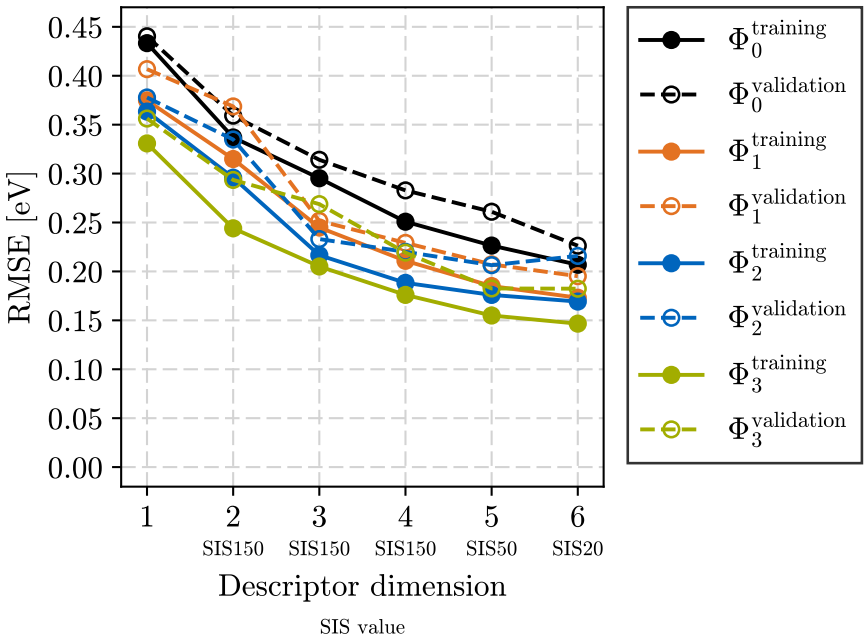


Figure S2: Training and validation results for a grid search of the two hyperparameters in the SISSO method (rung and dimension) using SIS values of 150 or the highest computationally tractable value for higher dimensions.

S3 Additional details on the screened alloy surfaces

Alloy phase diagrams can be constructed with the help of calculated thermodynamic descriptors like the bulk formation enthalpy H_f, which is available for a wide range of alloy structures in the AFLOW database.[9] Such a phase diagram for the case of the binary system Fe-Pt generated with the AFLOW-CHULL module is shown in Figure S3.[10] The envelope curtailing the different mixing ratios represents the convex hull, which is formed by the stable phases. The pure metals, thereby, serve as the reference with a formation enthalpy of zero. All other structures with an enthalpy between zero and the convex hull are metastable phases. The challenge is the variety of different space groups the structures belong to. In order to obtain reliable predictions we need to use structures that are similar to the training set, and we thereby need space groups that exhibit the same or a similar stepped facet. As the size of the supercell used for the DFT calculations limits the efficiency of the screening, only certain space groups are suitable for this purpose. These space groups need to preserve the symmetry of the alloy while not forming excessively large supercells. Note that due to alloying most space groups are not spherically symmetric. The atom arrangement of the facet depends on the orientation of the unit cell in three-dimensional space. Additionally, different atom arrangements of the same phase are possible depending on multiples of the chosen *hkl* plane. A careful analysis of the convex hull plots led to the choice of the space groups 221 and 225 (fcc structure) and 187 and 164 (hcp structure). Each of the fcc-like space groups has two different slab models, while the hcp-like groups are layered in the z-direction and show four (187 and 164) or two (187) different slab models depending on the TM ratio.

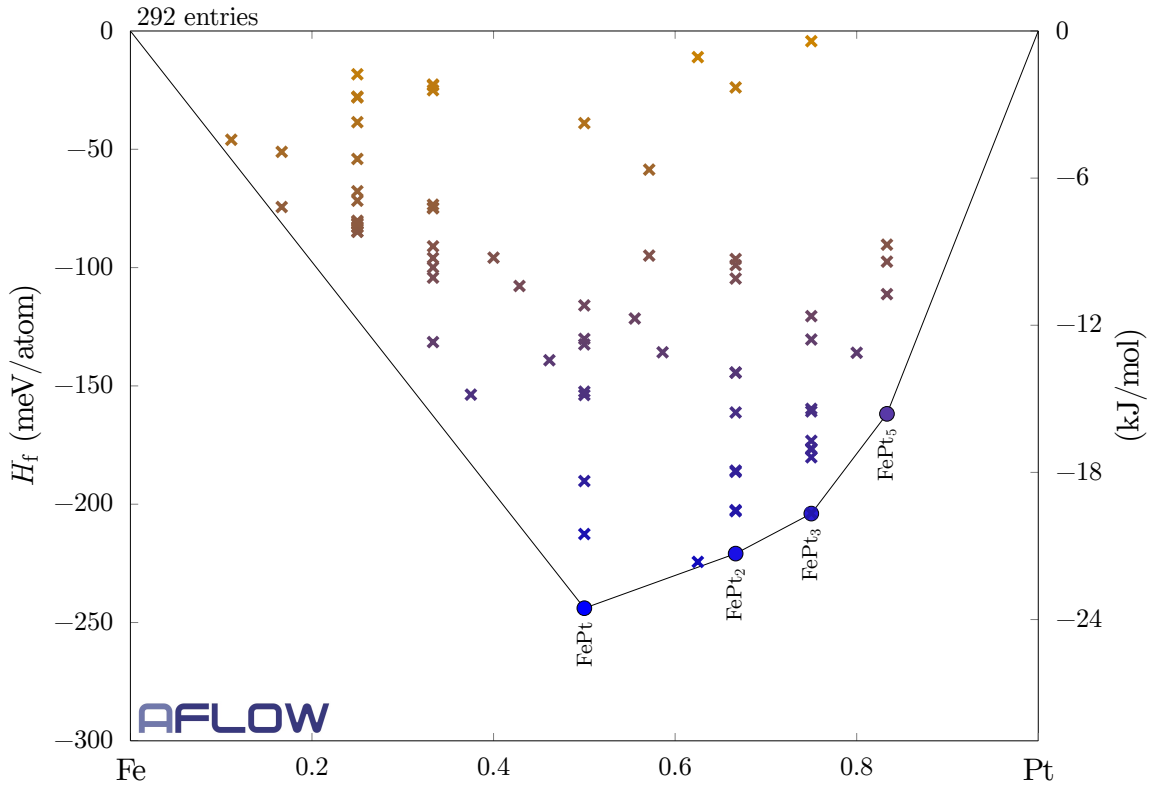


Figure S3: Example two-dimensional convex hull plot of Fe-Pt as generated by AFLOW-CHULL.[10] Phases located on the convex hull line (dots) are stable, while phases located between zero and the convex hull are metastable (crosses).

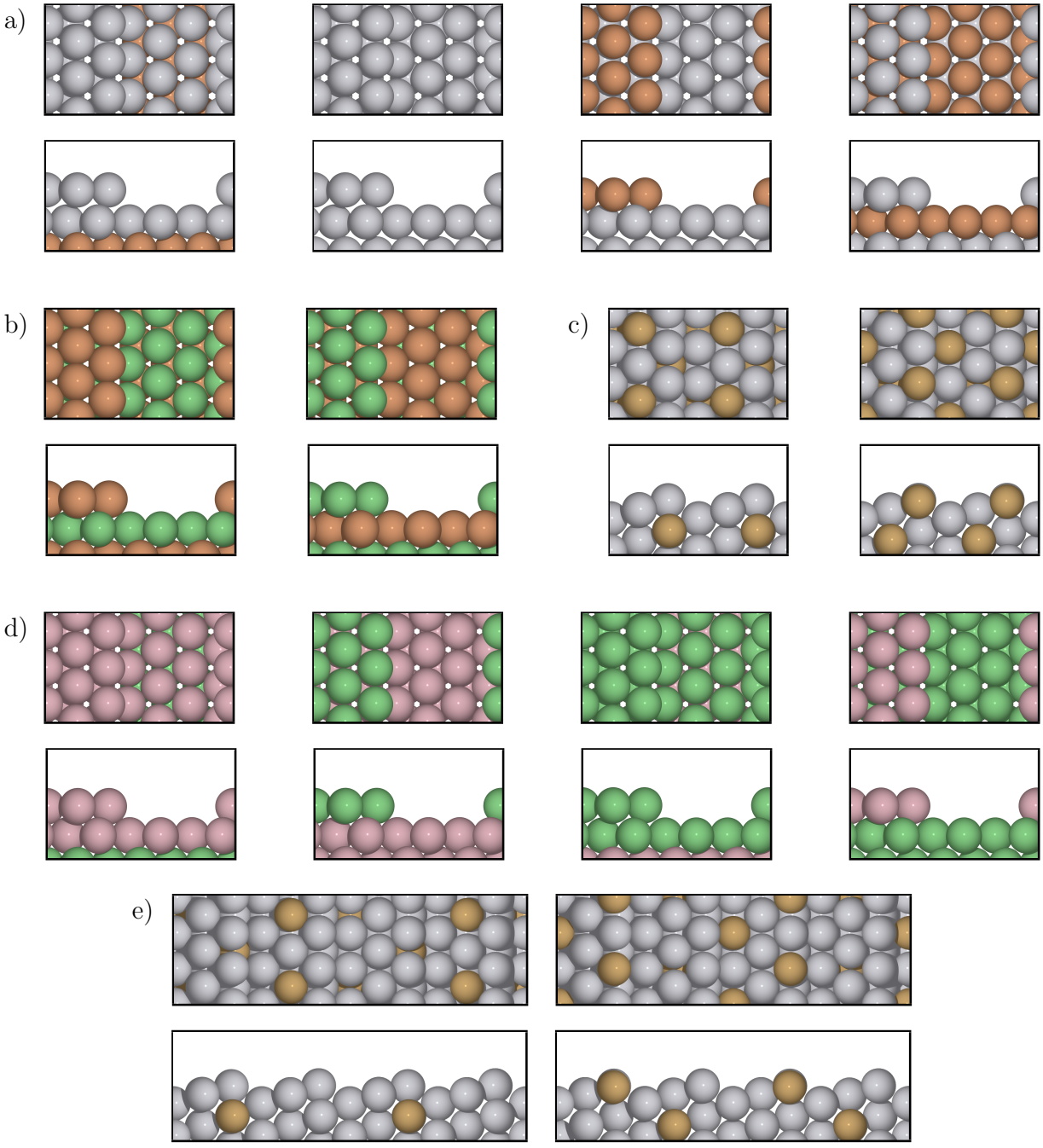


Figure S4: Depiction (top and side view) of the different alloy slab models. a) hcp alloy 3:1 SG187. b) hcp alloy 1:1 SG187. c) fcc alloy 3:1 SG221. d) hcp alloy 1:1 SG164. e) fcc alloy 7:1 SG225.

S4 Additional details on the compared microkinetic models

S4.1 Complex microkinetic model

The complex microkinetic model is taken from Andersen *et al.*[11] It consists of 24 reaction and 11 diffusion steps outlined in the following. It was found in ref 11 that within a mean-field model the barriers for the diffusion steps are not important. Therefore we adopted the same approximate diffusion barriers as used in the original work.

S4.1.1 Reaction Steps

$$\mathrm{H_{2(g)}} + 2\ *_\mathrm{h} \rightarrow 2\ \mathrm{H}^{*_\mathrm{h}} \tag{S1.1}$$

$$\mathrm{CO_{(g)}} + *_\mathrm{s} \rightarrow \mathrm{CO}^{*_\mathrm{s}} \tag{S1.2}$$

$$\mathrm{CO_{(g)}} + *_\mathrm{t} \rightarrow \mathrm{CO}^{*_\mathrm{t}} \tag{S1.3}$$

$$\mathrm{CO}^{*_\mathrm{s}} + \mathrm{H}^{*_\mathrm{h}} + *_\mathrm{f} + *_\mathrm{t} \rightleftharpoons \mathrm{C\text{-}OH}^{*_\mathrm{f}} + *_\mathrm{h} + *_\mathrm{s} + *_\mathrm{t} \rightarrow \mathrm{C}^{*_\mathrm{f}} + \mathrm{OH}^{*_\mathrm{t}} + *_\mathrm{h} + *_\mathrm{s} \tag{S1.4}$$

$$\mathrm{CO}^{*_\mathrm{t}} + \mathrm{H}^{*_\mathrm{h}} + *_\mathrm{f} \rightleftharpoons \mathrm{C\text{-}OH}^{*_\mathrm{f}} + *_\mathrm{h} + *_\mathrm{t} \rightarrow \mathrm{OH}^{*_\mathrm{t}} + \mathrm{C}^{*_\mathrm{f}} + *_\mathrm{h} \tag{S1.5}$$

$$\mathrm{C}^{*_\mathrm{f}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{C\text{-}H}^{*_\mathrm{f}} + *_\mathrm{h} \rightarrow \mathrm{CH}^{*_\mathrm{f}} + *_\mathrm{h} \tag{S1.6}$$

$$\mathrm{C}^{*_\mathrm{s}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{C\text{-}H}^{*_\mathrm{s}} + *_\mathrm{h} \rightarrow \mathrm{CH}^{*_\mathrm{s}} + *_\mathrm{h} \tag{S1.7}$$

$$\mathrm{C}^{*_\mathrm{t}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{C\text{-}H}^{*_\mathrm{t}} + *_\mathrm{h} \rightarrow \mathrm{CH}^{*_\mathrm{t}} + *_\mathrm{h} \tag{S1.8}$$

$$\mathrm{CH}^{*_\mathrm{f}} + \mathrm{H}^{*_\mathrm{h}} + *_\mathrm{t} \rightleftharpoons \mathrm{CH\text{-}H}^{*_\mathrm{f}} + *_\mathrm{h} + *_\mathrm{t} \rightarrow \mathrm{CH_2}^{*_\mathrm{t}} + *_\mathrm{f} + *_\mathrm{h} \tag{S1.9}$$

$$\mathrm{CH}^{*_\mathrm{f}} + \mathrm{H}^{*_\mathrm{h}} + *_\mathrm{s} \rightleftharpoons \mathrm{CH\text{-}H}^{*_\mathrm{f}} + *_\mathrm{h} + *_\mathrm{s} \rightarrow \mathrm{CH_2}^{*_\mathrm{s}} + *_\mathrm{f} + *_\mathrm{h} \tag{S1.10}$$

$$\mathrm{CH}^{*_\mathrm{s}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{CH\text{-}H}^{*_\mathrm{s}} + *_\mathrm{h} \rightarrow \mathrm{CH_2}^{*_\mathrm{s}} + *_\mathrm{h} \tag{S1.11}$$

$$\mathrm{CH}^{*_\mathrm{t}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{CH\text{-}H}^{*_\mathrm{t}} + *_\mathrm{h} \rightarrow \mathrm{CH_2}^{*_\mathrm{t}} + *_\mathrm{h} \tag{S1.12}$$

$$\mathrm{CH_2}^{*_\mathrm{s}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{CH_2\text{-}H}^{*_\mathrm{s}} + *_\mathrm{h} \rightarrow \mathrm{CH_3}^{*_\mathrm{s}} + *_\mathrm{h} \tag{S1.13}$$

$$\mathrm{CH_2}^{*_\mathrm{t}} + \mathrm{H_h} \rightleftharpoons \mathrm{CH_2\text{-}H}^{*_\mathrm{t}} + *_\mathrm{h} \rightarrow \mathrm{CH_3}^{*_\mathrm{t}} + *_\mathrm{h} \tag{S1.14}$$

$$CH_3^{*_s} + H^{*_h} \rightleftharpoons CH_3\text{-}H^{*_s} + *_h \qquad \rightarrow CH_{4(g)} + *_h + *_s \tag{S1.15}$$

$$CH_3^{*_t} + H^{*_h} \rightleftharpoons CH_3\text{-}H^{*_t} + *_h \qquad \rightarrow CH_{4(g)} + *_h + *_t \tag{S1.16}$$

$$OH^{*_s} + H^{*_h} \rightleftharpoons H\text{-}OH^{*_s} + *_h \qquad \rightarrow H_2O_{(g)} + *_h + *_s \tag{S1.17}$$

$$OH^{*_t} + H^{*_h} \rightleftharpoons H\text{-}OH^{*_t} + *_h \qquad \rightarrow H_2O_{(g)} + *_h + *_t \tag{S1.18}$$

$$2\ OH^{*_s} \rightleftharpoons H_2O\text{-}O^{*_s} + *_s \qquad \rightarrow H_2O_{(g)} + O^{*_s} + *_s \tag{S1.19}$$

$$2\ OH^{*_t} \rightleftharpoons H_2O\text{-}O^{*_t} + *_t \qquad \rightarrow H_2O_{(g)} + O^{*_t} + *_t \tag{S1.20}$$

$$OH^{*_s} + OH^{*_t} \rightleftharpoons H_2O\text{-}O\text{-}^{*_s} + *_t \qquad \rightarrow H_2O_{(g)} + O^{*_s} + *_t \tag{S1.21}$$

$$OH^{*_t} + OH^{*_s} \rightleftharpoons H_2O\text{-}O\text{-}^{*_t} + *_s \qquad \rightarrow H_2O_{(g)} + O^{*_t} + *_s \tag{S1.22}$$

$$O^{*_s} + H^{*_h} \rightleftharpoons O\text{-}H^{*_s} + *_h \qquad \rightarrow OH^{*_s} + *_h \tag{S1.23}$$

$$O^{*_t} + H^{*_h} \rightleftharpoons O\text{-}H^{*_t} + *_h \qquad \rightarrow OH^{*_t} + *_h \tag{S1.24}$$

S4.1.2 Diffusion Steps

$$CO^{*_s} + *_t \rightleftharpoons CO^{*_t} + *_s \tag{S1.25}$$

$$C^{*_f} + *_s \rightleftharpoons C^{*_s} + *_f \tag{S1.26}$$

$$C^{*_s} + *_t \rightleftharpoons C^{*_t} + *_s \tag{S1.27}$$

$$C^{*_t} + *_f \rightleftharpoons C^{*_f} + *_t \tag{S1.28}$$

$$CH^{*_f} + *_s \rightleftharpoons CH^{*_s} + *_f \tag{S1.29}$$

$$CH^{*_s} + *_t \rightleftharpoons CH^{*_t} + *_s \tag{S1.30}$$

$$CH^{*_t} + *_f \rightleftharpoons CH^{*_f} + *_t \tag{S1.31}$$

$$CH_2^{*_s} + *_t \rightleftharpoons CH_2^{*_t} + *_s \tag{S1.32}$$

$$CH_3^{*_s} + *_t \rightleftharpoons CH_3^{*_t} + *_s \tag{S1.33}$$

$$OH^{*_s} + *_t \rightleftharpoons OH^{*_t} + *_s \tag{S1.34}$$

$$O^{*_s} + *_t \rightleftharpoons O^{*_t} + *_s \tag{S1.35}$$

S4.2 Simple model

The simple microkinetic model is adopted from Lausche *et al.*[12] It is composed of 10 reaction steps.

S4.2.1 Reaction Steps

$$\mathrm{H_{2(g)} + 2\ *_h \rightarrow 2\ H^{*_h}} \tag{S2.1}$$

$$\mathrm{CO_{(g)} + *_s \rightarrow CO^{*_s}} \tag{S2.2}$$

$$\mathrm{CO^{*_s} + H^{*_h} + *_f \rightleftharpoons C\text{-}OH^{*_f} + *_s + *_h \rightarrow OH^{*_s} + C^{*_f} + *_h} \tag{S2.3}$$

$$\mathrm{C^{*_f} + H^{*_h} \rightleftharpoons C\text{-}H^{*_f} + *_h \rightarrow CH^{*_f} + *_h} \tag{S2.4}$$

$$\mathrm{CH^{*_f} + H^{*_h} + *_t \rightleftharpoons CH\text{-}H^{*_f} + *_h + *_t \rightarrow CH_2^{*_t} + *_f + *_h} \tag{S2.5}$$

$$\mathrm{CH_2^{*_t} + H^{*_h} \rightleftharpoons CH_2\text{-}H^{*_t} + *_h \rightarrow CH_3^{*_t} + *_h} \tag{S2.6}$$

$$\mathrm{CH_3^{*_t} + H^{*_h} \rightleftharpoons CH_3\text{-}H^{*_t} + *_h \rightarrow CH_{4(g)} + *_h + *_t} \tag{S2.7}$$

$$\mathrm{O^{*_s} + H^{*_h} \rightleftharpoons O\text{-}H^{*_s} + *_h \rightarrow OH^{*_s} + *_h} \tag{S2.8}$$

$$\mathrm{OH^{*_s} + H^{*_h} \rightleftharpoons H\text{-}OH^{*_s} + *_h \rightarrow H_2O_{(g)} + *_h + *_s} \tag{S2.9}$$

$$\mathrm{2\ OH^{*_s} \rightleftharpoons H_2O\text{-}O^{*_s} + *_s \rightarrow H_2O_{(g)} + O^{*_s} + *_s} \tag{S2.10}$$

S4.3 Reductionist Microkinetic Model

The reaction steps of the reductionist microkinetic model are the same as in the simple model (c.f. section S4.2). It differs from the latter only by which adsorption sites are considered. For each adsorbate this pool is restricted to the most stable site on Ni as determined with DFT. Consequently, steps S3.5–S3.7 differ from steps S2.5–S2.7 by the site group involved in CH_2 and CH_3 adsorption.

S4.3.1 Reaction Steps

$$\mathrm{H_{2(g)}} + 2\ *_\mathrm{h} \rightarrow 2\ \mathrm{H}^{*_\mathrm{h}} \tag{S3.1}$$

$$\mathrm{CO_{(g)}} + *_\mathrm{s} \rightarrow \mathrm{CO}^{*_\mathrm{s}} \tag{S3.2}$$

$$\mathrm{CO}^{*_\mathrm{s}} + \mathrm{H}^{*_\mathrm{h}} + *_\mathrm{f} \rightleftharpoons \mathrm{C\text{-}OH}^{*_\mathrm{f}} + *_\mathrm{s} + *_\mathrm{h} \rightarrow \mathrm{OH}^{*_\mathrm{s}} + \mathrm{C}^{*_\mathrm{f}} + *_\mathrm{h} \tag{S3.3}$$

$$\mathrm{C}^{*_\mathrm{f}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{C\text{-}H}^{*_\mathrm{f}} + *_\mathrm{h} \rightarrow \mathrm{CH}^{*_\mathrm{f}} + *_\mathrm{h} \tag{S3.4}$$

$$\mathrm{CH}^{*_\mathrm{f}} + \mathrm{H}^{*_\mathrm{h}} + *_\mathrm{s} \rightleftharpoons \mathrm{CH\text{-}H}^{*_\mathrm{f}} + *_\mathrm{h} + *_\mathrm{s} \rightarrow \mathrm{CH_2}^{*_\mathrm{s}} + *_\mathrm{f} + *_\mathrm{h} \tag{S3.5}$$

$$\mathrm{CH_2}^{*_\mathrm{s}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{CH_2\text{-}H}^{*_\mathrm{s}} + *_\mathrm{h} \rightarrow \mathrm{CH_3}^{*_\mathrm{s}} + *_\mathrm{h} \tag{S3.6}$$

$$\mathrm{CH_3}^{*_\mathrm{s}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{CH_3\text{-}H}^{*_\mathrm{s}} + *_\mathrm{h} \rightarrow \mathrm{CH_{4(g)}} + *_\mathrm{h} + *_\mathrm{s} \tag{S3.7}$$

$$\mathrm{O}^{*_\mathrm{s}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{O\text{-}H}^{*_\mathrm{s}} + *_\mathrm{h} \rightarrow \mathrm{OH}^{*_\mathrm{s}} + *_\mathrm{h} \tag{S3.8}$$

$$\mathrm{OH}^{*_\mathrm{s}} + \mathrm{H}^{*_\mathrm{h}} \rightleftharpoons \mathrm{H\text{-}OH}^{*_\mathrm{s}} + *_\mathrm{h} \rightarrow \mathrm{H_2O_{(g)}} + *_\mathrm{h} + *_\mathrm{s} \tag{S3.9}$$

$$2\ \mathrm{OH}^{*_\mathrm{s}} \rightleftharpoons \mathrm{H_2O\text{-}O}^{*_\mathrm{s}} + *_\mathrm{s} \rightarrow \mathrm{H_2O_{(g)}} + \mathrm{O}^{*_\mathrm{s}} + *_\mathrm{s} \tag{S3.10}$$

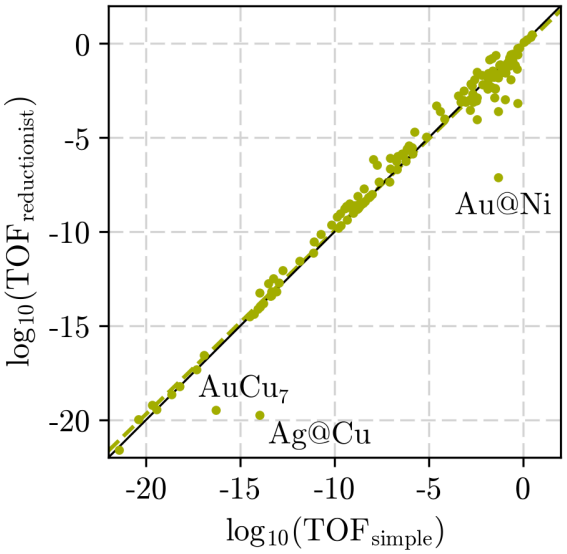


Figure S5: Correlation plot of the predicted TOFs obtained with the simple (x-axis) and reductionist (y-axis) microkinetic models.

S4.4 Degree of rate control

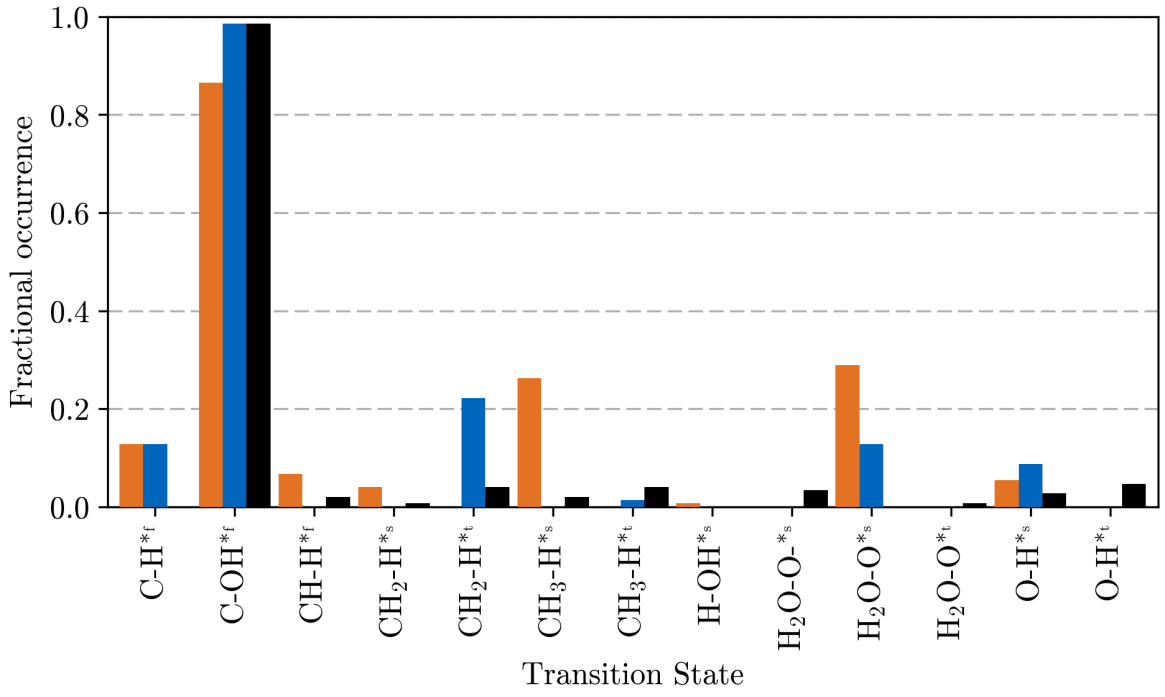


Figure S6: Fractional occurrence of a TS with an absolute DRC larger than 0.1 for the entire set of screened alloy surfaces, differentiating the compared microkinetic models (orange: reductionist, blue: simple, black: complex).

S4.5 Catalyst activities

Table S5: TOFs of those elemental TM catalysts and selected alloy surfaces that form the training data set, comparing the predictions from the three different microkinetic models. The adsorption energetics used for calculating the TOF is that predicted by the SISSO descriptor.

Catalyst	$\log_{10}(\mathrm{TOF}_{\mathrm{complex}})$	$\log_{10}(\mathrm{TOF}_{\mathrm{simple}})$	$\log_{10}(\mathrm{TOF}_{\mathrm{reductionist}})$
AgPd	−19.825	−18.661	−18.657
Cu	−15.899	−14.509	−14.509
Ag@Cu	−15.626	−13.993	−19.741
Pt	−13.154	−8.448	−7.709
PtRh	−11.910	−5.117	−4.964
Ir	−8.826	−1.604	−1.617
Pd	−8.756	−6.151	−5.578
IrRu	−8.532	−0.305	−0.579
Pd@Ir	−8.383	−1.528	−2.871
Pt@Rh	−8.031	−2.827	−3.527
Rh	−7.900	−2.482	−3.030
Ru	−5.440	−0.428	−1.155
Ni	−3.542	−1.305	−1.810
Au@Ni	−3.366	−1.328	−7.120
$\mathrm{Co}_{\mathrm{fcc}}$	−2.278	0.051	0.074
$\mathrm{Fe}_{\mathrm{bcc}}$	−0.892	−0.952	−2.968
$\mathrm{Co}_{\mathrm{hcp}}$	−0.752	0.214	0.198

Table S6: TOFs of those elemental TM catalysts and selected alloy surfaces that form the training data set, comparing the predictions from the three different microkinetic models. The adsorption energetics used for calculating the TOF is from the DFT training data (with a few highlighted exceptions).

Metal	$\log_{10}(\mathrm{TOF}^{\mathrm{DFT}}_{\mathrm{complex}})$	$\log_{10}(\mathrm{TOF}^{\mathrm{DFT}}_{\mathrm{simple}})$	$\log_{10}(\mathrm{TOF}^{\mathrm{DFT}}_{\mathrm{reductionist}})$
AgPd	−16.789	−13.301*	−13.372*
Cu	−16.363	−13.121	−13.121*
Ag@Cu	−15.823	−13.664	−18.570*
Pt	−12.275	−10.078	−7.621*
Pd@Ir	−9.696	−4.290	−3.575*’
Ir	−9.525	−3.597	−2.304*’
PtRh	−8.891	−4.695	−4.967
Pd	−8.530	−5.554	−5.846
IrRu	−5.067	0.792	−1.492
Rh	−4.137	−1.285	−3.343
Pt@Rh	−3.963	−1.421	−2.820*’
Ni	−2.769	−0.531	−2.253
Ru	−2.540	−0.231	−1.431
Au@Ni	−2.245	−0.653	−7.552*’
$\mathrm{Co}_{\mathrm{hcp}}$	−1.232	0.467	0.101
$\mathrm{Fe}_{\mathrm{bcc}}$	−1.085	−2.797*	−5.146*’
$\mathrm{Co}_{\mathrm{fcc}}$	−0.886	0.307	0.220

* using one or more energetics from the SISSO descriptor due to lack of appropriate DFT energetics for the model

’ sensitive to SISSO substitute energetics

References

(1) Andersen, M.; Levchenko, S. V.; Scheffler, M.; Reuter, K. Beyond Scaling Relations for the Description of Catalytic Materials. *ACS Catal.* **2019**, *9*, 2752–2759.

(2) Bengtsson, L. Dipole correction for surface supercell calculations. *Phys. Rev. B* **1999**, *59*, 12301–12304.

(3) Winter, M. J. WebElements. `https://www.webelements.com`, accessed: 2020-06-04.

(4) 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.

(5) Harrison, W. A. *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond*; Dover Publications, 1989.

(6) 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.

(7) Ouyang, R. SISSO: A data-driven method based on compressed-sensing for identifying descriptors for materials properties and functions. `https://github.com/rouyang2017/SISSO`, 2019; accessed: 2020-05-12.

(8) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

(9) Curtarolo, S.; Setyawan, W.; Hart, G. L.; Jahnatek, M.; Chepulskii, R. V.; Taylor, R. H.; Wang, S.; Xue, J.; Yang, K.; Levy, O.; Mehl, M. J.; Stokes, H. T.; Dem-

chenko, D. O.; Morgan, D. AFLOW: An automatic framework for high-throughput materials discovery. *Comput. Mater. Sci.* **2012**, *58*, 218–226.

(10) Oses, C.; Gossett, E.; Hicks, D.; Rose, F.; Mehl, M. J.; Perim, E.; Takeuchi, I.; Sanvito, S.; Scheffler, M.; Lederer, Y.; Levy, O.; Toher, C.; Curtarolo, S. AFLOW-CHULL: Cloud-Oriented Platform for Autonomous Phase Stability Analysis. *J. Chem. Inf. Model.* **2018**, *58*, 2477–2490.

(11) 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.

(12) Lausche, A. C.; Medford, A. J.; Khan, T. S.; Xu, Y.; Bligaard, T.; Abild-Pedersen, F.; Nørskov, J. K.; Studt, F. On the effect of coverage-dependent adsorbate-adsorbate interactions for CO methanation on transition metal surfaces. *J. Catal.* **2013**, *307*, 275–282.

Publication # 2

Selectivity Trends and Role of Adsorbate–Adsorbate Interactions in CO Hydrogenation on Rhodium Catalysts

Martin Deimel, Hector Prats, Michael Seibt, Karsten Reuter,
and Mie Andersen
ACS Catal. **12**, 7907 (2022).
DOI: 10.1021/acscatal.2c02353

ACS Catalysis

pubs.acs.org/acscatalysis

Research Article

Selectivity Trends and Role of Adsorbate−Adsorbate Interactions in CO Hydrogenation on Rhodium Catalysts

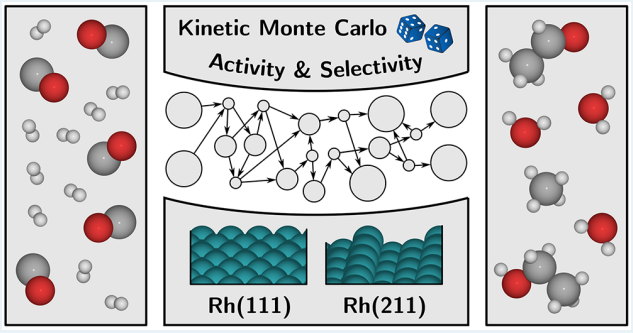
Martin Deimel, Hector Prats, Michael Seibt, Karsten Reuter, and Mie Andersen*

Cite This: *ACS Catal.* 2022, 12, 7907−7917

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Predictive-quality computational modeling of heterogeneously catalyzed reactions has emerged as an important tool for the analysis and assessment of activity and activity trends. In contrast, more subtle selectivities and selectivity trends still pose a significant challenge to prevalent microkinetic modeling approaches that typically employ a mean-field approximation (MFA). Here, we focus on CO hydrogenation on Rh catalysts with the possible products methane, acetaldehyde, ethanol, and water. This reaction has already been subjected to a number of experimental and theoretical studies with conflicting views on the factors controlling activity and selectivity toward the more valuable higher oxygenates. Using accelerated first-principles kinetic Monte Carlo simulations and explicitly and systematically accounting for adsorbate−adsorbate interactions through a cluster expansion approach, we model the reaction on the low-index Rh(111) and stepped Rh(211) surfaces. We find that the Rh(111) facet is selective toward methane, while the Rh(211) facet exhibits a similar selectivity toward methane and acetaldehyde. This is consistent with the experimental selectivity observed for larger, predominantly (111)-exposing Rh nanoparticles and resolves the discrepancy with earlier first-principles MFA microkinetic work that found the Rh(111) facet to be selective toward acetaldehyde. While the latter work tried to approximately account for lateral interactions through coverage-dependent rate expressions, our analysis demonstrates that this fails to sufficiently capture concomitant correlations among the adsorbed reaction intermediates that crucially determine the overall selectivity.



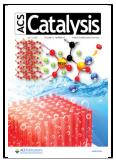
KEYWORDS: *heterogeneous catalysis, CO hydrogenation, microkinetic modeling, kinetic Monte Carlo, density functional theory*

INTRODUCTION

The conversion of syngas (CO and H_2) into hydrocarbons and oxygenates is attractive as an alternative source of fuels and chemicals. However, selectivity toward the more useful higher oxygenates such as ethanol and acetaldehyde remains challenging, with methane being a common undesired product.[1−4] Many theoretical and experimental works have focused on understanding and tuning especially the selectivity of Rh catalysts, as Rh is generally recognized as one of the most promising elemental catalysts for the direct synthesis of higher oxygenates. It has now become clear that pure Rh catalysts are intrinsically selective toward primarily methane and acetaldehyde,[5,6] whereas ethanol synthesis requires promoters such as Fe and Mn.[5,7−11]

Recent experimental works have suggested that there is an inverse relationship between activity and selectivity for pure Rh catalysts, where an overall higher activity (CO conversion) correlates with a lower selectivity toward acetaldehyde. Explanations offered for this trend are, however, conflicting. Yang et al. have suggested, on the basis of density functional theory (DFT) and mean-field microkinetics, that it is the nature of the active sites exposed by the catalyst nanoparticles that is the deciding factor, with step sites being highly active and selective toward methane and terrace sites being less active and selective toward acetaldehyde.[5] This view was recently challenged by Schumann et al.[6] They synthesized Rh nanoparticles of different sizes and found that it is primarily small particles below 2 nm that exhibit high acetaldehyde selectivity and low activity, while the larger particles above 5 nm are the most active and selective toward methane. The surface fraction of edge/corner sites increases for smaller particles and was found to closely follow the selectivity trends. While this could indicate that step or corner sites are the active sites for acetaldehyde synthesis, Schumann et al. proposed instead that the smaller particles support a much higher local CO coverage at both terrace and step sites, which limits the

Received: May 12, 2022
Revised: June 2, 2022
Published: June 17, 2022



https://doi.org/10.1021/acscatal.2c02353
ACS Catal. 2022, 12, 7907−7917

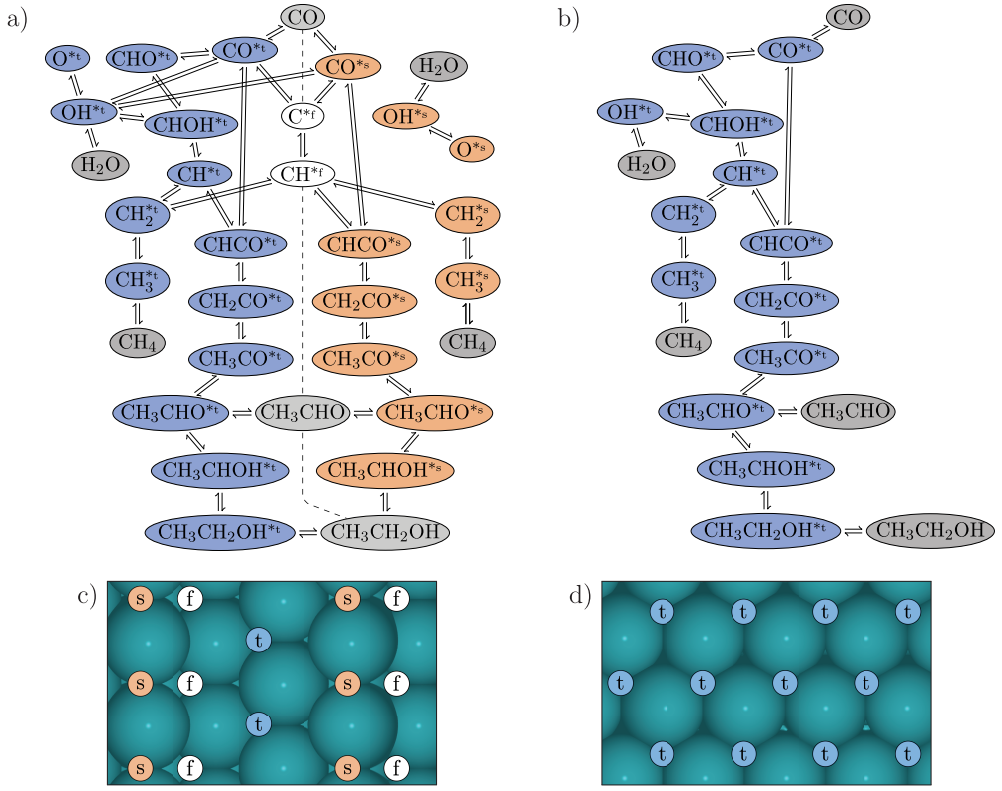


Figure 1. Reaction networks of (a) the Rh(211) and (b) the Rh(111) KMC models, excluding diffusion steps (these connect e.g., OH^{*_t} and OH^{*_s}). For hydrogenation reactions an additional $H^{*_{S/T}}$ species is implied. On the right side of the dashed line in (a) are shown the elementary steps occurring on the s sites (orange) and on the left side are shown the corresponding reactions on the t sites (blue). Adsorbates on f sites are shown in white and gaseous molecules are shown in gray. Images of the corresponding Rh(211) and Rh(111) facets with the respective sites are shown in (c) and (d), respectively.

activity due to poisoning but increases the acetaldehyde selectivity by driving C + CO coupling reactions.

The conflicting views offered in the literature clearly call for more thorough theoretical investigations and improved microkinetic models that can account for the effects of high surface coverages and concomitantly increased lateral adsorbate−adsorbate interactions. In the mean-field study carried out by Yang et al., coverage-dependent rate equations were parametrized and employed to mimic such interactions. However, it is well recognized in the literature that mean-field kinetics cannot properly account for the effects of correlations and fluctuations, including fluctuations in the local coverage.[12−15] A more accurate approach is kinetic Monte Carlo (KMC) simulations using a cluster expansion (CE) to treat lateral interactions.[16−20] Unfortunately, this can become very expensive for complex lateral interaction models and large disparities in the time scales of the different processes. In a recent work by Chen et al. on syngas conversion at the Rh(111) surface,[21] some of these challenges were avoided by applying their hybrid XPK method,[22] which is based on the assumption that there is a clear time scale separation between fast diffusion processes and slow reactions such that the system can be solved by alternating between separate KMC and mean-field models. By comparing the hybrid method to a static mean-field model parametrized to include lateral interactions, Chen et al. were able to show that the dynamic account of fluctuations in the local coverage achieved in the XPK model is crucial to correctly capture experimental selectivity trends, where higher total pressures have been shown to lead to higher acetaldehyde selectivity.[21] Their assumption that diffusion limitations do not cause any spatial correlations in the distribution of adsorbates on the surface may, however, not always be applicable, as we will show in this work and has also been demonstrated in previous literature works.[12,23]

In this work we revisit the question of the role played by step and terrace sites for activity and selectivity trends in syngas conversion over Rh catalysts. We consider the pristine Rh(211) and Rh(111) facets, which are representative of step and terrace sites found at larger nanoparticles where finite-size effects no longer play a large role. Due to recent methodological developments in our in-house KMC code kmos[24] concerning the efficient modeling of lateral interactions and the implementation of an acceleration algorithm[25,26] to tackle the time-scale disparity problem, we are able to carry out full-blown KMC simulations employing a CE model for lateral interactions. The results with and without an account of lateral interactions are compared to the corresponding mean-field kinetics.

The main finding of our work is that—in contrast to previously parametrized coverage-dependent mean-field models—KMC simulations that explicitly and systematically account for lateral interactions are able to correctly capture the experimental selectivity trends for large nanoparticles. We show and rationalize why in some cases the lateral interactions have a huge effect on the results (Rh(111) facet), whereas in other cases the effects are negligible (Rh(211) facet). The

https://doi.org/10.1021/acscatal.2c02353
ACS Catal. 2022, 12, 7907−7917

finding of a breakdown of mean-field kinetics is not restricted to models with lateral interactions. In fact, we show that, also in the absence of lateral interactions, reaction-induced inhomogeneities and diffusion limitations can cause the mean-field-predicted activities to deviate substantially from the KMC results (Rh(211) facet). Finally, we show that, in order to reach a quantitative agreement with both the selectivity and the activity trends observed in experiments, we need to correct for well-known errors in DFT-predicted adsorption energies such as the persistent overbinding of CO with standard semilocal DFT functionals.[27−29]

METHODS

Reaction Model and Parametrization. The reaction networks employed for the Rh(211) and Rh(111) facets are shown in Figure 1 and were inspired by the work of Yang et al.[5] As in our previous works on CO hydrogenation over stepped metal surfaces,[26,30] and extending over the work of Yang et al., we use a highly resolved active site representation for the Rh(211) facet consisting of a terrace site t, an upper step site s, and a lower 4-fold coordinated step site f. Only the last two sites were considered by Yang et al. The full reaction networks we employed can be found in Section S1 of the Supporting Information, together with a discussion of why other reaction steps considered previously in the literature[1,2,31] are less plausible than those considered here. DFT data calculated at both low and high CO coverages were taken from Yang et al. and used to parametrize a CE model, which was used in connection with the KMC simulations. Some additional DFT calculations were carried out by us using the Quantum Espresso code[32] with the BEEF-vdW functional[33] and the exact same numerical settings as used by Yang et al. (see Section S2 of the Supporting Information).

Adsorption energies are referenced to CO, H_2O and CH_3OH using the formation energy approach

$$E^x_{\mathrm{form}} = E^x_{\mathrm{slab+ads}} - E_{\mathrm{slab}} - \sum_{i\in x} n^x_i \mu_i$$

with the formation energy E^x_{form} of adsorbate x, the total energy of the adsorbed species on the slab $E^x_{\mathrm{slab+ads}}$, that of the clean slab E_{slab}, and the gas-phase reference energy μ_i of atom i as well as the occurrence n^x_i of atom i in the adsorbate x. Adsorbate−adsorbate interactions are incorporated using a CE that is terminated after the first term corresponding to pairwise nearest neighbor interactions

$$E^M_{\mathrm{form}} = \sum_{x\in M} n^M_x E^x_{\mathrm{form}} + \sum_j k^M_j \varepsilon_j$$

with the occurrence n^M_x and formation energies E^x_{form} of the individual adsorbates x in structure M and the occurrence k^M_j and interaction energy ε_j of the pairwise interaction j. The interaction energies (cf. Table S2 in the Supporting Information) are obtained by solving a set of linear equations for a number of different structures M. As the surface under the investigated conditions is mostly covered with CO, we consider only self-interactions between the CO molecules, interactions between CO and the remaining adsorbates (except H), and selected other interactions (e.g., between C and CH at the f site) following Yang et al. All other interactions are neglected due to their low probability of occurrence. Following common practice,[5,34] H is adsorbed at special "hydrogen reservoir" sites, which reflects the assumption that it can intercalate into sublattice sites because of its smaller size in comparison to the other adsorbates. In the KMC models, we implement this by treating H in a mean-field ansatz, following our previous work.[26] This means that it is not an actual species but is only represented via an effective H coverage. We use different energetics (and thereby coverage) depending on whether the H atom is involved in reactions taking place at an s or a t site (for H we denote these reservoir sites as S and T sites in both the KMC and mean-field models). Our treatment thereby implicitly assumes that there are no spatial correlations in the distribution of H at the surface, which should be a good approximation when it is taken into account that H diffusion and adsorption/desorption are fast processes. Note, however, that the assumption that H does not block any surface sites or interact with the other adsorbates would most likely break down under high-pressure reaction conditions.

For the rate constants of elementary steps corresponding to adsorption, desorption, reaction, and diffusion, we used standard expressions, including zero-point energies and other enthalpy/entropy corrections within the harmonic approximation (adsorbates and transition states) or the ideal gas approximation (gas species) from the thermochemistry module of the Atomic Simulation Environment software package[35,36] (cf. our previous work[26]). Vibrational frequencies are taken from Yang et al.[5] In our previous work,[26] diffusion steps of CO, O, OH, CH, CH_2, and CH_3 were found to not be rate-limiting and therefore we stick to approximate barriers calculated for Re(0001) from Hahn et al.[37] For same-site diffusion of C at f sites and CH at t sites we use values calculated for Rh(211) from our previous work.[26] Diffusion of all other species is neglected. To allow for comparison of our simulations to those of Yang et al.,[5] we use the same reaction conditions of p_{CO} = 13.33 bar, $p_{\mathrm{H_2}}$ = 6.66 bar, and $p_{\mathrm{H_2O}} = p_{\mathrm{CH_4}} = p_{\mathrm{CH_3CHO}} = p_{\mathrm{CH_3CH_2OH}}$ = 0 bar at the three different temperatures 523, 585, and 650 K.

In order to model the effect of lateral interactions on the reaction kinetics during the simulations, the energy barriers E^a of the elementary steps for the possible lattice configurations are linearly approximated using Brønsted−Evans−Polanyi (BEP) relations[38,39]

$$E^a = \alpha(\Delta E_{\mathrm{FS}} - \Delta E_{\mathrm{IS}}) + E^a_0$$

with the energy shift ΔE_{FS} (ΔE_{IS}) of the final (initial) state due to lateral interactions, the zero-coverage barrier E^a_0, and the parameter α with values in the interval [0, 1] representing a reactant-like (0) or product-like (1) transition state. DFT-calculated energetics for determining the α parameters shown in Table S3 are taken from Yang et al. and Andersen et al.[5,26] Although BEP relations entail some energetic uncertainty, the magnitude is on the order of the error introduced by semilocal DFT.[16] The approach of combining CE with BEP relations is used across different KMC frameworks and effectively reduces the computational burden for the inclusion of lateral interactions. Previous studies employing this approach were able to quantitatively capture experimental observations without the need to explicitly calculate all possible lattice configurations, which is especially critical with an increasing number of sites and adsorbates in the reaction network under consideration.[16,40−42]

Kinetic Monte Carlo. The KMC simulation technique allows for a numerical solution to the time evolution of the spatial distribution of the adsorbates on the coarse-grained

https://doi.org/10.1021/acscatal.2c02353
ACS Catal. 2022, 12, 7907−7917

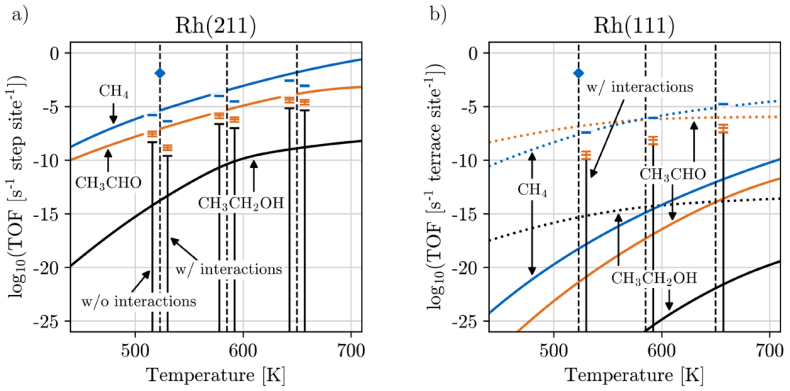


Figure 2. TOFs as a function of the temperature for the (a) Rh(211) and (b) Rh(111) facets for CH_4 (blue), CH_3CHO (orange), and CH_3CH_2OH (black). The MFA results without (with) lateral interactions are shown with solid lines (dotted lines), and the KMC results are shown at 523, 585, and 650 K (dashed vertical lines) without (left offset) and with (right offset) lateral interactions. The blue diamonds at 523 K represent the experimentally measured total TOF (primarily methane) of the largest nanoparticles with diameters above 5 nm from Schumann et al.[6]

sites of the catalytic surface, along with related properties such as the occurrence of individual reaction pathways and the catalytic activity and selectivity in terms of the turnover frequencies (TOFs) for the formation of the reaction products. Since individual elementary steps are executed step by step with probabilities reflected by their rate constants, KMC can suffer from a performance bottleneck in the case of processes that occur on disparate time scales: e.g., fast diffusion and slow reactions. This time scale disparity challenge is tackled here through an acceleration algorithm developed by Dybeck et al.[25] In a recent work[26] we implemented this algorithm in the kmos code[24] and used it to study trends in CO methanation activity over stepped transition metals. While some challenging cases were observed, the algorithm was found to work well for the mainly CO-covered Rh(211) facet. Our previous work disregarded lateral interactions, however, and these can play an important role in the outcome of a reaction, as shown in this work. Both repulsive and attractive interactions with neighboring adsorbates can alter the energetics of elementary steps and lead to changes in their individual rates. In this work we used the acceleration algorithm in connection with the recently developed on-the-fly backend in kmos.[43,44] This backend features increased performance and reduced memory requirements for models with many lateral interactions in comparison to the original backend, since the rates of processes affected by lateral interactions are calculated at runtime according to the aforementioned CE model and BEP relations.

The error bars for the TOFs are obtained from a Bayesian error analysis. Reactions in KMC follow the Poisson distribution

$$P(n|\nu) = \frac{e^{-\nu}\nu^n}{n!}$$

where $P(n|\nu)$ is the probability for observing n turnovers during a fixed simulation time t, given the expected value ν. Here, we are rather interested in calculating the probability distribution for ν, given a (possibly small or even zero) number of observed turnover events n. This posterior distribution can be obtained via Bayes' theorem

$$P(\nu|n) = \frac{P(n|\nu)P_0(\nu)}{\int_0^\infty P(n|\nu)P_0(\nu)\,d\nu} = \frac{e^{-\nu}\nu^n P_0(\nu)}{\int_0^\infty e^{-\nu}\nu^n P_0(\nu)\,d\nu}$$

with the prior probability distribution $P_0(\nu)$, which we assume to be a constant C as there is no *a priori* information on the TOF. To normalize $P_0(\nu)$, we set ν_{max}, i.e., $C = 1/\nu_{max}$, to a value high enough that the following approximation holds:

$$\frac{e^{-\nu}\nu^n}{\int_0^{\nu_{max}} e^{-\nu}\nu^n\,d\nu} \approx \frac{e^{-\nu}\nu^n}{\int_0^\infty e^{-\nu}\nu^n\,d\nu} = \frac{e^{-\nu}\nu^n}{n!}$$

This gives the posterior probability distribution

$$P(\nu|n) = \frac{e^{-\nu}\nu^n}{n!}$$

which is also a Poisson distribution. Note that this also gives us the probability distribution for the TOF, since the TOF is simply ν divided by the simulation time t. The most likely value of the TOF is determined from the maximum of the posterior distribution, and the upper and lower bounds of the error bars are obtained as the smallest credible interval that contains 99% of the total probability mass.

Sensitivity analysis was carried out by calculating the degree of rate control (DRC) proposed by Campbell and co-workers[45] for each TS i ($X_{RC,i}$)

$$X_{RC,i} = \left(\frac{\partial \ln r}{\partial\left(\frac{-G_i}{RT}\right)}\right)_{G_{j\neq i}}$$

with the rate r, the free energy G_i of TS i, the universal gas constant R, and the absolute temperature T. In KMC the derivative was approximated by the finite-difference expression

$$X_{RC,i} = \left(\frac{\ln r_+ - \ln r_-}{\frac{-0.2\ \text{eV}}{RT}}\right)_{G_{j\neq i}}$$ with r_+ (r_-) being the rate for an increase (decrease) of the TS energy by 0.1 eV. For the determination of the associated error bars, the respective upper ν_+ and lower bounds ν_- of the 99% credible interval for the two energy modifications are combined to contain ν_+ of one and ν_- of the other simulation and *vice versa*. The upper (lower) limit of the selectivities of methane, acetaldehyde, and ethanol are obtained by considering the ν_+ (ν_-) value of the 99% credible interval of one product and the respective ν_- (ν_+) values of the other two products.

The pair probabilities of second-order processes are obtained by storing the time-integrated counts of lattice

https://doi.org/10.1021/acscatal.2c02353
ACS Catal. 2022, 12, 7907−7917

configurations, in which the process can be executed over the entire simulation, and the counts are then normalized by the total simulation time and the size of the system. All simulations were run for 5×10^7 steps to reach a steady state and subsequently until 11 (Rh(111)) or 26 turnovers (Rh(211)) of acetaldehyde were observed or the total KMC simulation time reached 1 week. The simulations were carried out using a 10×10 lattice for Rh(211) and a 25×25 lattice for Rh(111) with periodic boundary conditions. Further details about the KMC simulation settings and the convergence tests carried out for both the lattice size as well as the parameters employed in the acceleration algorithm can be found in Section S3 of the Supporting Information.

Mean-Field Microkinetic Modeling. For comparison to the KMC simulations, we also carried out microkinetic simulations using the CatMAP software package.[46] CatMAP employs the mean-field approximation (MFA), meaning that the spatial distribution of the adsorbates is further coarse-grained into a mean coverage of each site type, thereby also neglecting coverage fluctuations around the mean. This results in a set of coupled rate equations, which are solved at the steady state. All of our MFA simulations were performed without lateral interactions, as the parametrization of coverage-dependent rate equations is not a topic of this work. We emphasize that the MFA simulations used exactly the same reaction network and adsorption energetics as our KMC simulations, so that differences between both simulation approaches are entirely due to the MFA employed.

RESULTS AND DISCUSSION

We begin by presenting MFA and KMC results for the Rh(211) facet in Figure 2a. The KMC results are reported at three different temperatures without (left offset) and with (right offset) lateral interactions. Without lateral interactions, a comparison of our MFA results (solid lines) and KMC results (left offset) reveals that the activities are slightly overestimated in the MFA model by about a factor of 2.7−6.0 and that both models show similar selectivity trends toward methane rather than acetaldehyde. Ethanol turnovers are not observed in the KMC simulations—hence the large error bars—which is consistent with the very low TOFs obtained in the MFA model. We will come back to the reason for the differences in actual TOF prediction between MFA and KMC below.

Turning to the effect of lateral interactions, we can observe from the KMC results that they do not influence the TOFs much. Our MFA results are only presented without lateral interactions, as the parametrization of these are not available for our modified active site representation. However, in the original Rh(211) MFA model from Yang et al. (see Figure S1 in the Supporting Information), lateral interactions do influence the TOFs, especially at lower temperatures. At 523 K the difference ranges from a factor of about 300 for ethanol to about 4000 for acetaldehyde.

Before diving into a deeper analysis of the differences between the different models for the Rh(211) facet, we present also the TOFs for the Rh(111) facet (see Figure 2b). This facet contains only a terrace site, and thus the active site representation and reaction network we employ (cf. Section S1.3 in the Supporting Information) is in this case completely identical with that of Yang et al. Comparing the MFA results with and without interactions (dotted and solid lines), we can see that for this facet the influence of the lateral interactions is huge—at all temperatures the catalyst is essentially inactive if lateral interactions are not taken into account. Furthermore, it is surprising to note that the lateral interactions in the MFA model, which were parametrized using a second-order expansion in the coverage, give rise to a change in the selectivity trends. That is, without interactions the catalyst is always selective toward methane, but with interactions the catalyst is selective toward acetaldehyde at the experimentally relevant temperature below about 600 K. We recall that the MFA results for Rh(211) and Rh(111) with interactions were used by Yang et al. to propose that the activity−selectivity trends obtained with experimentally synthesized Rh nanoparticles are caused by varying amounts of step and terrace sites at these particles, where step (terrace) sites would then be the active sites for methane (acetaldehyde) formation, respectively.[5]

However, as discussed in the Introduction, this explanation was recently challenged by Schumann et al.[6] on the basis of new and more detailed experiments which showed that large particles above 5 nm, which are expected to predominantly expose the Rh(111) facet, are actually selective toward methane and not acetaldehyde. Interestingly, we can fully confirm this from our KMC simulations with lateral interactions, where at all simulated temperatures we find that the Rh(111) facet is indeed selective toward methane (see Figure 2b). Without interactions we do not observe any meaningful TOFs for any of the products (not shown), which is consistent with the very low TOFs obtained in the MFA model without interactions. Thus, our KMC results confirm that lateral interactions are of paramount importance for the Rh(111) facet but the CE used to parametrize these interactions in KMC gives results qualitatively different from those of the MFA-parametrized lateral interaction model. Importantly, only the explicit site-resolving KMC results are able to reproduce the experimentally observed selectivity trends for large particles. The case of small particles is outside the scope of the present work, as the low-index Rh(211) and Rh(111) surfaces mainly represent step and terrace sites at larger particles free from finite-size effects.

We believe that the shortcomings of the MFA models for both Rh(211) and Rh(111) are caused by the well-known problems of these models with accounting for effects of correlations and fluctuations, including fluctuations in the local coverage.[12−15] In the following we exemplarily analyze this for the Rh(211) surface without interactions. This analysis starts by performing a sensitivity analysis to identify the rate-limiting steps (RLSs) of the two models. As shown in Figure 3, the RLS for both the methane and the acetaldehyde pathways is mainly water formation at the t site in the KMC model, whereas for the MFA model it is mainly methane formation and CH−CO coupling at the s site. The sensitivity analysis for the KMC model with interactions (see Figure S3 in the Supporting Information) is very similar to that of the KMC model without interactions.

Since a breakdown of the MFA is typically associated with second-order reaction steps where correlations in the spatial distribution of the two reacting species at the surface occur,[26] we next evaluate the pair probabilities for selected elementary steps from our KMC simulations. In Figure 4 we plot the ratio between the KMC-simulated pair probability to find the reacting species A and B at neighboring sites and the MFA-assumed probability equal to $c[A][B]$, where c is a geometric factor (the site connectivity) and [A] ([B]) is the surface coverage of species A (B). It is seen that the MFA indeed

https://doi.org/10.1021/acscatal.2c02353
ACS Catal. 2022, 12, 7907−7917

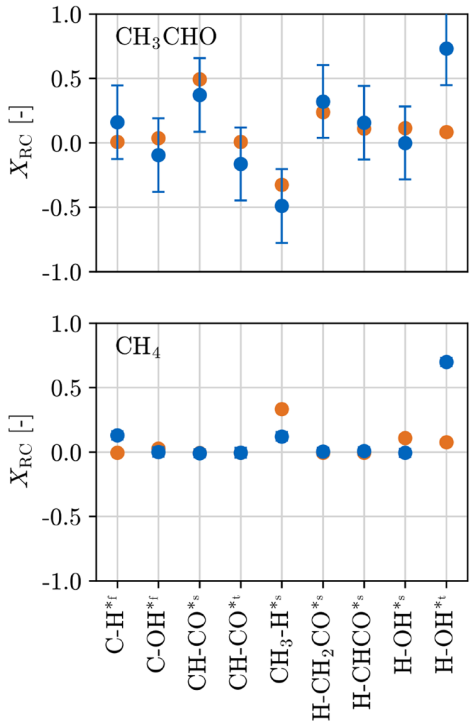


Figure 3. Sensitivity analysis of the KMC (blue points) and MFA (orange points) models for Rh(211) without lateral interactions, showing the DRC for CH_4(g) (bottom) and CH_3CHO(g) (top) of selected transition states. The analysis is performed at 650 K.

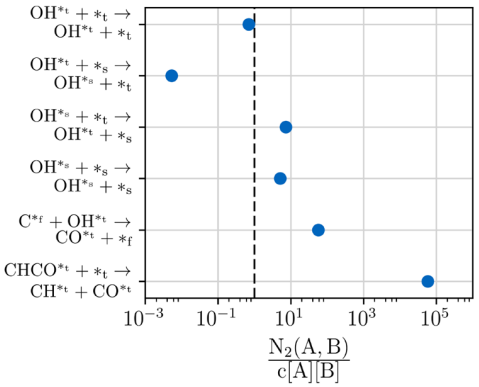


Figure 4. Ratio between KMC and MFA probabilities to find the pair of reacting species in selected second-order reaction steps at neighboring sites in the Rh(211) model without lateral interactions. The analysis is performed at 650 K.

breaks down for several of these steps: i.e., the ratio is significantly different from 1. In particular, we note that the MFA breaks down for the reaction between C at the f site and OH at the t site to form CO at the t site and H. This is the reverse of the CO dissociation step, which is executed just before the RLS in the KMC model, where OH at the t site is further hydrogenated to form water. The probability of finding the two reactant species, C^{*_f} and OH^{*_t}, next to each other is about 57 times larger in KMC than in MFA. This is caused by the high barrier for C to diffuse along the step and the high CO coverage at the t site, which effectively hinders OH diffusion at the t sites (cf. our previous work on the comparison of MFA and KMC models for CO hydrogenation at stepped metals[26]).

As a consequence of this persistent correlation, it becomes more probable for C^{*_f} and OH^{*_t} to react back to form CO^{*_t} and H in the KMC model in comparison to the MFA model. In a free energy diagram (cf. orange and black lines in Figure 5a), this change in pair probability can be represented as an increase in the free energy barrier for the reverse reaction step of $C^{*_f} + OH^{*_t}$ in the MFA model (orange line) caused by a decrease in the free energy of the $C^{*_f} + OH^{*_t}$ state (i.e., a higher configurational entropy in the MFA-assumed well-mixed state). Since the free energy barrier for further reaction out of the $C^{*_f} + OH^{*_t}$ state (i.e., water formation at the t site) is unchanged, the effective barrier in the MFA free energy landscape (difference between $H{-}OH^{*_t}$ transition state and CO^{*_t} state) is smaller than the effective barrier in the KMC free energy landscape. This explains the larger TOF obtained in the MFA model and the change of the RLS from water formation at the t site to methane formation and CH−CO coupling at the s site (the latter steps are not shown in Figure 5a).

From the free energy diagram in Figure 5a we can also explain why the KMC TOF and RLS for both methane and acetaldehyde formation at the Rh(211) facet does not change much upon inclusion of lateral interactions. As seen by a comparison of the black curve (without interactions) and the blue curve (with interactions), the free energies of the states that determine the effective barrier ($H{-}OH^{*_t}$ transition state and CO^{*_t} state) move up in energy by a similar amount as a consequence of interactions, causing the effective barrier and TOF to remain at a similar value. The detailed lateral interaction parameters at play in the two states are illustrated in Figure S4 in the Supporting Information.

While they are unimportant for the Rh(211) model, we already mentioned above that lateral interactions do have a huge effect on the Rh(111) model, and we will next analyze why this is the case. The free energy diagram with and without lateral interactions for the Rh(111) facet is shown in Figure 5b. The greatest difference in comparison to the Rh(211) facet is that here lateral interactions actually push the free energy of the CO^{*_t} state up close to the free energy of gas-phase CO. Thereby, the average CO coverage decreases from around 100% without interactions to about 62% with interactions at the analyzed temperature of 650 K (cf. Section S3.3 in the Supporting Information). In Figure 5b the CO^{*_t} state is split into two levels, since CO desorption primarily happens from a local high-coverage state where on average 5.1 out of 6 neighboring sites are occupied by CO, whereas hydrogenation of CO primarily happens from a lower-coverage state where only 3 neighboring sites are occupied by CO. The average coverage pattern with about $\frac{2}{3}$ CO coverage resembles a honeycomb lattice of CO molecules, as illustrated in Figure S5 in the Supporting Information. The RLS for both methane and acetaldehyde formation in the model with interactions is primarily dissociation of CHOH to form CH and OH (and for acetaldehyde also the CH_2CO hydrogenation step) (cf. Figure S6 in the Supporting Information).

For Rh(111), the profound effect of lateral interactions on the simulation results cannot be explained by a change to the effective free energy barrier (i.e., the difference between the

https://doi.org/10.1021/acscatal.2c02353
ACS Catal. 2022, 12, 7907−7917

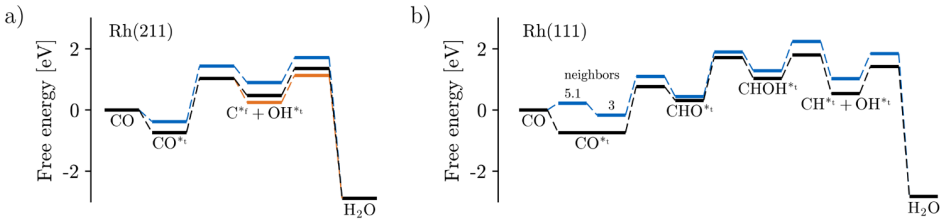


Figure 5. Free energy diagrams for the CO dissociation and water formation pathways in the (a) Rh(211) and (b) Rh(111) KMC model without (black) and with (blue) lateral interactions. In (a) the orange curve represents the modified probability for the reverse CO activation step in the MFA model without interactions (see text). The energy levels with lateral interactions for Rh(211) and CO desorption for Rh(111) are determined by the average barriers of the up to 10000 last executed events of each process. For the remaining levels at Rh(111) we do not have enough statistics: e.g., the back-reaction of CH and OH to form CHOH never occurs. Since these steps are mostly executed from the surface configuration illustrated in Figure S5, where each CO has three neighboring CO molecules, we instead determined the energy levels and barriers specifically for this surface configuration. The analyses are performed at 650 K.

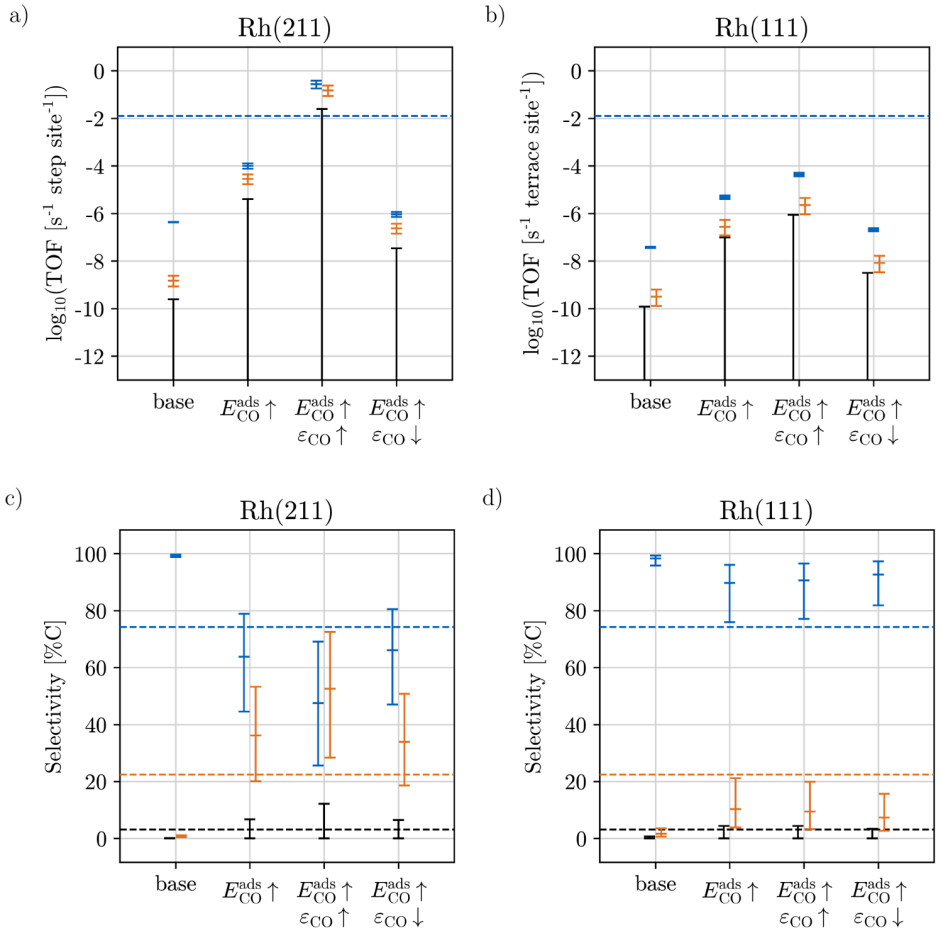


Figure 6. Variations in predicted KMC TOFs for the (a) Rh(211) and (b) Rh(111) facets and corresponding carbon selectivities for the (c) Rh(211) and (d) Rh(111) facets with modifications to the DFT parameters in the models with lateral interactions (see text). Results are shown for CH_4 (blue), CH_3CHO (orange), and CH_3CH_2OH (black). "base" represents the unmodified results also shown in Figure 2a,b. The dashed horizontal lines represent the experimentally measured total TOF (primarily methane, (a) and (b)) and selectivities ((c) and (d)) of the largest nanoparticles with diameters above 5 nm from Schumann et al.[6] Both the theoretical and the experimental results are obtained at 523 K.

rate-limiting CH−OH^{*_t} TS and the CO^{*_t} state in Figure 5b), since this barrier decreases with only around 132 meV upon inclusion of lateral interactions. Rather, the important point is that, for CHOH to dissociate, a neighboring t vacancy is required. Without lateral interactions, the surface is completely poisoned with an average CO coverage of close to 100%. Thus, the probability for CHOH dissociation to occur is negligible, which explains why no product formation could be observed in the KMC simulations. With interactions, however, a formed CHOH will on average have 3 CO and 3 vacancy neighbors (see Figure S5 in the Supporting Information), which makes the dissociation reaction possible. For the Rh(211) facet a similar CO poisoning is not observed, since here CO dissociation relies on a vacant f site, and the vacancy coverage

https://doi.org/10.1021/acscatal.2c02353
ACS Catal. 2022, 12, 7907−7917

of the f sites remains high even without lateral interactions (cf. Table S4).

Up until now, we have shown that the MFA can break down both in the absence of lateral interactions (analyzed for the Rh(211) facet where the TOFs are overestimated in the MFA model compared to KMC) and in the presence of lateral interactions (analyzed for the Rh(111) facet where wrong selectivity trends are obtained in the MFA model in comparison to experiments). However, even if the KMC model does recover the experimental selectivity trends, the absolute TOFs predicted by both the KMC model and the MFA model are not in good agreement with the experiments.

In Figure 2a,b we mark with a blue diamond the total TOF (primarily methane) measured for the largest Rh nanoparticles investigated by Schumann et al.[6] It is larger by about 5 orders of magnitude in comparison to the Rh(211) KMC simulation with lateral interactions (cf. Figure 2a) and by about 5−6 orders of magnitude for the Rh(111) KMC and MFA simulations with lateral interactions (cf. Figure 2b). The consistent, gross underestimation of the TOF by both KMC and MFA and for both investigated facets leads us to suspect that this discrepancy has its origin in the DFT calculations used to parametrize the kinetic models. In particular, it is well-known that generalized gradient approximation (GGA) functionals tend to overbind CO on transition metals:[29] e.g., the CO adsorption energy obtained with the BEEF-vdW functional employed here for Rh(111) (−1.7 eV) is about 0.25 eV stronger than the experimental value (−1.45 eV[28,47,48]). Since the state with CO adsorbed at a terrace site directly influences the effective barrier of the reaction (cf. Figure 5), this error directly influences the obtained TOF.

In order to assess the possible implications of this error, we show in Figure 6 the TOFs and selectivities obtained for the Rh(211) and Rh(111) facets for both the hitherto discussed base model and a model where the CO adsorption energies are increased by 0.25 eV ($E_{CO}^{ads}\uparrow$). This already improves the agreement with experiments significantly: i.e., the TOF predicted for the Rh(211) facet (Rh(111) facet) is now only about 2 (3) orders of magnitude lower than that of the experiment. Of course, this analysis neglects the fact that there could also be errors associated with other rate-controlling parameters in our kinetic model: e.g., the transition state for water formation at the terrace site or lateral interaction parameters affecting either the latter transition state or the CO adsorption energy. Since we do not know in which direction such other errors might point, we show as examples in Figure 6a,b also results where the interaction parameter for self-interactions between the CO molecules at t and s sites are increased ($\varepsilon_{CO}\uparrow$) or decreased ($\varepsilon_{CO}\downarrow$) by 50 meV from their base values. An error of this magnitude seems quite reasonable, since we use an approximate CE that is terminated after pairwise nearest-neighbor interactions. Taking these results as plausible "error bars" on the theoretical results, we can now reconcile theory with experiment to the extent that the experimental TOF lies within (admittedly rough) error bars on the theoretical values for the Rh(211) facet. In reality, the experimentally used nanoparticles of course contain both step and terrace sites, and the actual TOF measured might be the result of an interplay between reaction steps taking place at both site types. Such bifunctional effects have been previously discussed in the literature[49,50] and have indeed recently been demonstrated theoretically for e.g., CO oxidation at Pt nanoparticles[51] and the hydrogen evolution reaction at jagged Pt nanowires.[52] The finding here that the TOF is largest, and in best agreement with experiment, on the Rh(211) facet is related to the facile hydrogen-assisted CO dissociation at the step sites exposed by this facet.[31] However, the subsequent rate-controlling water formation step takes place at the terrace site; thus, both types of sites are required for a high activity. We would here again like to emphasize that the modeling carried out in this work is only applicable to larger nanoparticles that can be reliably represented through extended bulk-cut facets. In order to describe small particles where finite-size effects play a large role, realistic nanoparticle geometries would have to be employed both at the DFT and KMC level. This is computationally very challenging with currently available methods but would be an interesting addition for future work in order to understand also the experimentally observed activity and selectivity trends for small particles.

Comparing the experimental selectivities shown as dashed horizontal lines in Figure 6c,d with the theoretical selectivities reveals that for both facets variations in the adsorption energies and interaction energies (as well as combinations) bring the selectivities closer to the experimental measurements. The observed trend of an increase in the acetaldehyde selectivity with a destabilization of adsorbed CO for both Rh(211) and Rh(111) is caused by a decrease in the effective barriers in the acetaldehyde pathways involving the RLSs of CH−CO coupling (for Rh(211) as shown in Figure 3) and CH_2CO hydrogenation (for Rh(111) as shown in Figure S6). Nevertheless, for all parameter modifications, our results robustly indicate that the Rh(111) facet is selective to methane. For the Rh(211) facet the selectivity trends fluctuate between methane selectivity or a similar methane and acetaldehyde selecitivity, depending on the exact parameter values. Given that the 5 nm nanoparticles studied by Schumann et al. will dominantly expose the Rh(111) facet, as well as step and corner sites to a smaller extent, the sum of the theoretical selectivity trends should overall result in a methane selectivity in quite good agreement with the experimental measurements.

CONCLUSIONS

We investigated the CO hydrogenation reaction on Rh(111) and Rh(211) using accelerated first-principles KMC simulations with and without lateral interactions parametrized from a cluster expansion model. The results are compared to MFA simulations from Yang et al.[5] with and without coverage dependence in the rate expressions. The coverage-dependent MFA model predicts acetaldehyde selectivity for the Rh(111) facet below ∼580 K, which is in contrast to the KMC simulations with lateral interactions that predict methane selectivity at all temperatures. Importantly, only the KMC results are in agreement with recent detailed experimental investigations on selectivity trends of Rh nanoparticles from Schumann et al.[6]

The inclusion of lateral interactions is found to have a huge effect on the Rh(111) simulations. We explain this by the fact that the RLS in this model (CHOH dissociation) requires a neighboring vacant terrace site, which is a probable lattice configuration only when the reduction of the surface CO coverage caused by repulsive CO−CO interactions is taken into account. In contrast, for the Rh(211) model we find that the TOF and RLS (water formation at the terrace site) are only weakly affected by lateral interactions. This is explained by

https://doi.org/10.1021/acscatal.2c02353
ACS Catal. 2022, 12, 7907−7917

similar surface coverages with and without lateral interactions and by the fact that interactions shift the two states controlling the effective barrier in the dominant reaction pathway by similar amounts, leading to net similar effective barriers with and without interactions. For Rh(211) we furthermore show that the MFA also breaks down in the absence of interactions due to diffusion limitations and reaction-induced lattice inhomogeneities. This leads to higher TOFs and changes in the RLSs in the MFA model in comparison to the KMC model.

Finally, we compare the overall activities (methane and acetaldehyde TOFs) to the experimental measurements from Schumann et al. We find that we need to correct for the well-known CO overbinding of GGA functionals (about 0.25 eV for the here applied BEEF-vdW functional) in order to approach agreement with the experiments. This correction also improves the quantitative agreement with the carbon selectivities obtained by Schumann et al. In particular, it leads to a larger increase in the acetaldehyde TOFs in comparison to the methane TOFs for both Rh(211) and Rh(111), since the effective barrier of the acetaldehyde pathway involves CO both in the CO activation step leading to CH_x species at the surface (i.e., the step shared with the methane pathway) and in the subsequent step where the formed CH species reacts with another CO.

Overall, the insights obtained in this work could be relevant for further tailoring heterogeneous catalysts to improve their selectivity toward the desired higher oxygenates acetaldehyde and ethanol. The methodological advances demonstrated here, i.e., the combination of acceleration algorithms in KMC with an efficient modeling of lateral interactions, open up possibilities for also treating complex reaction networks at a level of detail beyond the hitherto applied approximate MFA models. This is important in order to achieve reliable mechanistic insights as a solid basis for the rational design of selective catalysts.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.2c02353.

Additional details of the compared microkinetic models, DFT, and KMC (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Mie Andersen − *Aarhus Institute of Advanced Studies, Aarhus University, 8000 Aarhus C, Denmark; Center for Interstellar Catalysis, Department of Physics and Astronomy, Aarhus University, 8000 Aarhus C, Denmark;* orcid.org/0000-0002-9943-1534; Email: mie@phys.au.dk

Authors

Martin Deimel − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, 85747 Garching, Germany;* orcid.org/0000-0001-7178-4832

Hector Prats − *Department of Chemical Engineering, University College London, London WC1E 7JE, UK;* orcid.org/0000-0003-4991-253X

Michael Seibt − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, 85747 Garching, Germany*

Karsten Reuter − *Fritz-Haber-Institut der Max-Planck-Gesellschaft, 14195 Berlin, Germany;* orcid.org/0000-0001-8473-8659

Complete contact information is available at: https://pubs.acs.org/10.1021/acscatal.2c02353

Notes

The authors declare no competing financial interest.

Input files for the kmos simulations for the different models together with additionally calculated DFT geometries are available at https://github.com/m-deimel/CO_hydrogenation.git.

■ ACKNOWLEDGMENTS

We acknowledge funding and support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2089/1-390776260. M.A. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 754513, the Aarhus University Research Foundation, the Danish National Research Foundation through the Center of Excellence "InterCat" (grant agreement No. DNRF150), and VILLUM FONDEN (grant No. 37381).

■ REFERENCES

(1) Choi, Y.; Liu, P. Mechanism of Ethanol Synthesis from Syngas on Rh(111). *J. Am. Chem. Soc.* **2009**, *131*, 13054−13061.

(2) 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.

(3) Luk, H. T.; Mondelli, C.; Ferré, D. C.; Stewart, J. A.; Pérez-Ramírez, J. Status and prospects in higher alcohols synthesis from syngas. *Chem. Soc. Rev.* **2017**, *46*, 1358−1426.

(4) Ao, M.; Pham, G. H.; Sunarso, J.; Tade, M. O.; Liu, S. Active Centers of Catalysts for Higher Alcohol Synthesis from Syngas: A Review. *ACS Catal.* **2018**, *8*, 7025−7050.

(5) Yang, N.; Medford, A. J.; Liu, X.; Studt, F.; Bligaard, T.; Bent, S. F.; Nørskov, J. K. Intrinsic Selectivity and Structure Sensitivity of Rhodium Catalysts for C_{2+} Oxygenate Production. *J. Am. Chem. Soc.* **2016**, *138*, 3705−3714.

(6) Schumann, M.; Nielsen, M. R.; Smitshuysen, T. E. L.; Hansen, T. W.; Damsgaard, C. D.; Yang, A.-C. A.; Cargnello, M.; Grunwaldt, J.-D.; Jensen, A. D.; Christensen, J. M. Rationalizing an Unexpected Structure Sensitivity in Heterogeneous Catalysis − CO Hydrogenation over Rh as a Case Study. *ACS Catal.* **2021**, *11*, 5189−5201.

(7) Hu, J.; Wang, Y.; Cao, C.; Elliott, D. C.; Stevens, D. J.; White, J. F. Conversion of biomass-derived syngas to alcohols and C_2 oxygenates using supported Rh catalysts in a microchannel reactor. *Catal. Today* **2007**, *120*, 90−95.

(8) Mei, D.; Rousseau, R.; Kathmann, S. M.; Glezakou, V.-A.; Engelhard, M. H.; Jiang, W.; Wang, C.; Gerber, M. A.; White, J. F.; Stevens, D. J. Ethanol synthesis from syngas over Rh-based/SiO_2 catalysts: A combined experimental and theoretical modeling study. *J. Catal.* **2010**, *271*, 325−342.

(9) Liu, Y.; Göeltl, F.; Ro, I.; Ball, M. R.; Sener, C.; Aragão, I. B.; Zanchet, D.; Huber, G. W.; Mavrikakis, M.; Dumesic, J. A. Synthesis Gas Conversion over Rh-Based Catalysts Promoted by Fe and Mn. *ACS Catal.* **2017**, *7*, 4550−4563.

(10) Yang, N.; Yoo, J. S.; Schumann, J.; Bothra, P.; Singh, J. A.; Valle, E.; Abild-Pedersen, F.; Nørskov, J. K.; Bent, S. F. Rh-MnO

Interface Sites Formed by Atomic Layer Deposition Promote Syngas Conversion to Higher Oxygenates. *ACS Catal.* **2017**, *7*, 5746−5757.
(11) Huang, X.; Teschner, D.; Dimitrakopoulou, M.; Fedorov, A.; Frank, B.; Kraehnert, R.; Rosowski, F.; Kaiser, H.; Schunk, S.; Kuretschka, C.; Schlögl, R.; Willinger, M.-G.; Trunschke, A. Atomic-Scale Observation of the Metal−Promoter Interaction in Rh-Based Syngas-Upgrading Catalysts. *Ang. Chem. Int. Ed.* **2019**, *58*, 8709−8713.
(12) Matera, S.; Meskine, H.; Reuter, K. Adlayer inhomogeneity without lateral interactions: Rationalizing correlation effects in CO oxidation at RuO_2(110) with first-principles kinetic Monte Carlo. *J. Chem. Phys.* **2011**, *134*, 064713.
(13) Stamatakis, M.; Piccinin, S. Rationalizing the Relation between Adlayer Structure and Observed Kinetics in Catalysis. *ACS Catal.* **2016**, *6*, 2105−2111.
(14) Liu, D.-J.; Zahariev, F.; Gordon, M. S.; Evans, J. W. Predictive Beyond-Mean-Field Rate Equations for Multisite Lattice−Gas Models of Catalytic Surface Reactions: CO Oxidation on Pd(100). *J. Phys. Chem. C* **2016**, *120*, 28639−28653.
(15) Li, X.; Grabow, L. C. Evaluating the benefits of kinetic Monte Carlo and microkinetic modeling for catalyst design studies in the presence of lateral interactions. *Catal. Today* **2022**, *387*, 150−158.
(16) Stamatakis, M. Kinetic modelling of heterogeneous catalytic systems. *J. Phys.: Condens. Matter* **2015**, *27*, 013001.
(17) Piccinin, S.; Stamatakis, M. Steady-State CO Oxidation on Pd(111): First-Principles Kinetic Monte Carlo Simulations and Microkinetic Analysis. *Top. Catal.* **2017**, *60*, 141−151.
(18) Jørgensen, M.; Grönbeck, H. Scaling Relations and Kinetic Monte Carlo Simulations To Bridge the Materials Gap in Heterogeneous Catalysis. *ACS Catal.* **2017**, *7*, 5054−5061.
(19) Andersen, M.; Panosetti, C.; Reuter, K. A Practical Guide to Surface Kinetic Monte Carlo Simulations. *Front. Chem.* **2019**, *7*, 202.
(20) Reuter, K. Ab Initio Thermodynamics and First-Principles Microkinetics for Surface Catalysis. *Catal. Lett.* **2016**, *146*, 541−563.
(21) Chen, Z.; Wang, H.; Liu, Z.; Xu, X. Dynamic and Intermediate-Specific Local Coverage Controls the Syngas Conversion on Rh(111) Surfaces: An Operando Theoretical Analysis. *ACS Catal.* **2021**, *11*, 3830−3841.
(22) Chen, Z.; Wang, H.; Su, N. Q.; Duan, S.; Shen, T.; Xu, X. Beyond Mean-Field Microkinetics: Toward Accurate and Efficient Theoretical Modeling in Heterogeneous Catalysis. *ACS Catal.* **2018**, *8*, 5816−5826.
(23) Temel, B.; Meskine, H.; Reuter, K.; Scheffler, M.; Metiu, H. Does phenomenological kinetics provide an adequate description of heterogeneous catalytic reactions? *J. Chem. Phys.* **2007**, *126*, 204711.
(24) Hoffmann, M. J.; Matera, S.; Reuter, K. kmos: A lattice kinetic Monte Carlo framework. *Comput. Phys. Commun.* **2014**, *185*, 2138−2150.
(25) Dybeck, E. C.; Plaisance, C. P.; Neurock, M. Generalized Temporal Acceleration Scheme for Kinetic Monte Carlo Simulations of Surface Catalytic Processes by Scaling the Rates of Fast Reactions. *J. Chem. Theory Comput.* **2017**, *13*, 1525−1538.
(26) 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.
(27) Feibelman, P. J.; Hammer, B.; Nørskov, J. K.; Wagner, F.; Scheffler, M.; Stumpf, R.; Watwe, R.; Dumesic, J. The CO/Pt(111) Puzzle. *J. Phys. Chem. B* **2001**, *105*, 4018−4025.
(28) Abild-Pedersen, F.; Andersson, M. CO adsorption energies on metals with correction for high coordination adsorption sites − A density functional study. *Surf. Sci.* **2007**, *601*, 1747−1753.
(29) Patra, A.; Peng, H.; Sun, J.; Perdew, J. P. Rethinking CO adsorption on transition-metal surfaces: Effect of density-driven self-interaction errors. *Phys. Rev. B* **2019**, *100*, 035442.
(30) Deimel, M.; Reuter, K.; Andersen, M. Active Site Representation in First-Principles Microkinetic Models: Data-Enhanced Computational Screening for Improved Methanation Catalysts. *ACS Catal.* **2020**, *10*, 13729−13736.
(31) Ferrin, P.; Simonetti, D.; Kandoi, S.; Kunkes, E.; Dumesic, J. A.; Nørskov, J. K.; Mavrikakis, M. Modeling Ethanol Decomposition on Transition Metals: A Combined Application of Scaling and Brønsted-Evans-Polanyi Relations. *J. Am. Chem. Soc.* **2009**, *131*, 5809−5815.
(32) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Corso, A. D.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.
(33) 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* **2012**, *85*, 235149.
(34) Lausche, A. C.; Medford, A. J.; Khan, T. S.; Xu, Y.; Bligaard, T.; Abild-Pedersen, F.; Nørskov, J. K.; Studt, F. On the effect of coverage-dependent adsorbate−adsorbate interactions for CO methanation on transition metal surfaces. *J. Catal.* **2013**, *307*, 275−282.
(35) Bahn, S. R.; Jacobsen, K. W. An object-oriented scripting interface to a legacy electronic structure code. *Comput. Sci. Eng.* **2002**, *4*, 56−66.
(36) Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; Hermes, E. D.; Jennings, P. C.; Jensen, P. B.; Kermode, J.; Kitchin, J. R.; Kolsbjerg, E. L.; Kubal, J.; Kaasbjerg, K.; Lysgaard, S.; Maronsson, J. B.; Maxson, T.; Olsen, T.; Pastewka, L.; Peterson, A.; Rostgaard, C.; Schiøtz, J.; Schütt, O.; Strange, M.; Thygesen, K. S.; Vegge, T.; Vilhelmsen, L.; Walter, M.; Zeng, Z.; Jacobsen, K. W. The atomic simulation environment − a Python library for working with atoms. *J. Phys.: Condens. Matter* **2017**, *29*, 273002.
(37) Hahn, K.; Mavrikakis, M. Atomic and Molecular Adsorption on Re(0001). *Top. Catal.* **2014**, *57*, 54−68.
(38) Brønsted, J. N. Acid and Basic Catalysis. *Chem. Rev.* **1928**, *5*, 231−338.
(39) Evans, M. G.; Polanyi, M. Inertia and driving force of chemical reactions. *Trans. Faraday Soc.* **1938**, *34*, 11−24.
(40) Wu, C.; Schmidt, D.; Wolverton, C.; Schneider, W. Accurate coverage-dependence incorporated into first-principles kinetic models: Catalytic NO oxidation on Pt (111). *J. Catal.* **2012**, *286*, 88−94.
(41) Nielsen, J.; d'Avezac, M.; Hetherington, J.; Stamatakis, M. Parallel kinetic Monte Carlo simulation framework incorporating accurate models of adsorbate lateral interactions. *J. Chem. Phys.* **2013**, *139*, 224706.
(42) Piccinin, S.; Stamatakis, M. CO Oxidation on Pd(111): A First-Principles-Based Kinetic Monte Carlo Study. *ACS Catal.* **2014**, *4*, 2143−2152.
(43) Lorenzi, J. M. *Ab initio modeling of heterogeneous catalysis: Multiscale analysis of interacting reaction pathways*. Ph.D. thesis, Technical University of Munich, 2017.
(44) Seibt, M. *Temporal Acceleration of Kinetic Monte Carlo Simulations Applied to Heterogeneous Catalysis*. M.Sc. thesis, Technical University of Munich, 2018.
(45) Stegelmann, C.; Andreasen, A.; Campbell, C. T. Degree of Rate Control: How Much the Energies of Intermediates and Transition States Control Rates. *J. Am. Chem. Soc.* **2009**, *131*, 8077−8082.
(46) 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.
(47) Wei, D.; Skelton, D.; Kevan, S. Desorption and molecular interactions on surfaces: CO/Rh(110), CO/Rh(100) and CO/Rh(111). *Surf. Sci.* **1997**, *381*, 49−64.
(48) Wellendorff, J.; Silbaugh, T. L.; Garcia-Pintos, D.; Nørskov, J. K.; Bligaard, T.; Studt, F.; Campbell, C. T. A benchmark database for

https://doi.org/10.1021/acscatal.2c02353
ACS Catal. 2022, 12, 7907−7917

adsorption bond energies to transition metal surfaces and comparison to selected DFT functionals. *Surf. Sci.* **2015**, *640*, 36−44.
(49) Andersen, M.; Medford, A. J.; Nørskov, J. K.; Reuter, K. Analyzing the Case for Bifunctional Catalysis. *Angew. Chem., Int. Ed.* **2016**, *55*, 5210−5214.
(50) Andersen, M.; Medford, A. J.; Nørskov, J. K.; Reuter, K. Scaling-Relation-Based Analysis of Bifunctional Catalysis: The Case for Homogeneous Bimetallic Alloys. *ACS Catal.* **2017**, *7*, 3960−3967.
(51) Jørgensen, M.; Grönbeck, H. The Site-Assembly Determines Catalytic Activity of Nanoparticles. *Angew. Chem., Int. Ed.* **2018**, *57*, 5086−5089.
(52) Gu, G. H.; Lim, J.; Wan, C.; Cheng, T.; Pu, H.; Kim, S.; Noh, J.; Choi, C.; Kim, J.; Goddard, W. A.; Duan, X.; Jung, Y. Autobifunctional Mechanism of Jagged Pt Nanowires for Hydrogen Evolution Kinetics via End-to-End Simulation. *J. Am. Chem. Soc.* **2021**, *143*, 5355−5363.

https://doi.org/10.1021/acscatal.2c02353
ACS Catal. 2022, 12, 7907−7917

Supporting Information: Selectivity Trends and Role of Adsorbate-Adsorbate Interactions in CO Hydrogenation on Rhodium Catalysts

Martin Deimel,[†] Hector Prats,[‡] Michael Seibt,[†] Karsten Reuter,[¶] and Mie Andersen[*,§,‖]

[†]*Chair for Theoretical Chemistry and Catalysis Research Center, Technical University of Munich, Lichtenbergstr. 4, 85747 Garching, Germany*

[‡]*Department of Chemical Engineering, University College London, Roberts Building, Torrington Place, London WC1E 7JE, UK*

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

[§]*Aarhus Institute of Advanced Studies, Aarhus University, Aarhus C, DK-8000 Denmark*

[‖]*Department of Physics and Astronomy - Center for Interstellar Catalysis, Aarhus University, Aarhus C, DK-8000 Denmark*

E-mail: mie@phys.au.dk

S1 Additional Details on the Microkinetic Models

S1.1 KMC Reaction Network on Rh(211)

The KMC Rh(211) model is adopted from Yang *et al.*[1] and extended by expanding their step site into a separate step (s) and terrace (t) site. It is composed of 34 reaction and 19 diffusion steps. For this network, the C-C coupling is assumed to only occur *via* the reaction of CH+CO either on the s or on t sites. Other C-C coupling reactions were not included for the following reasons:

- **C-CO coupling to CCO**

  On s sites, this step has an energy barrier of 1.16 eV.[2] However, it is very endothermic, leading to an energy barrier for the reverse direction of only 0.28 eV.[2] This cannot compete with the C hydrogenation reaction to produce CH, which has an energy barrier of 1.04 eV[1] and is almost thermoneutral. To our knowledge, this step has never been considered on t sites.

- **CH_2-CO coupling to CH_2CO**

  On s sites, this step has an energy barrier of 1.07 eV.[2] However, this would hardly compete with CH_2 hydrogenation to CH_3, which has an energy barrier of only 0.42 eV.[1] On t sites, this step has been reported to be highly activated,[3] with a barrier of 1.34 eV,[4] which cannot compete with CH_2 hydrogenation.

- **CH_3-CO coupling to CH_3CO**

  On s sites, this step has an energy barrier of 1.38 eV. However, this would hardly compete with CH_3 hydrogenation to methane, which has an energy barrier of only 0.55 eV.[1] On t sites, this step was included by Choi *et al.*[4] in their mechanistic study, finding this process cannot compete with the production of CH_4 unless its energy barrier is significantly decreased.

Regarding CO hydrogenation to CHO or COH, both possibilities have been considered for the

Rh(211) facet in the present study (see steps 1.20 and 1.3/1.4, respectively, in Section S1.1.1 below). Note that, in the case of COH, the COH formation step and subsequent dissociation step have been coarse-grained into a single elementary step ($CO + H \rightleftharpoons C + OH$) to speed up the KMC simulation. This is expected to be a good approximation, since the COH intermediate is very high in energy and thus expected to be short-lived compared to the initial and final states in step 1.3/1.4. The energy of the transition state is taken as that of the COH dissociation, since it is the highest in energy. This approximation was already used in a previous study.[5] For the Rh(111) facet (see Section S1.3.1 below) we include only the CO dissociation pathway going through the intermediates CHO, CHOH and finally dissociating to CH and OH. The reason is that Yang *et al.*[1] found that for Rh(111) this pathway has an overall lower barrier at high CO coverage conditions than the pathway going through the COH intermediate and dissociating to C and OH, which is the dominant pathway for Rh(211).

Finally, another pathway that was not included here is the hydrogenation of CHO to CH_2O, CH_3O, and CH_3OH. The reason for this choice is that the selectivity of Rh catalysts towards methanol is very low.[1] Therefore, we believe all the above-mentioned alternative reaction pathways are less plausible than those investigated here, and that we can extract solid mechanistic understanding from our simplified reaction model, where comparison with experimental results can be done.

S1.1.1 Reaction Steps

$$\mathrm{CO_{(g)}} + *_\mathrm{s} \rightleftharpoons \mathrm{CO}^{*_\mathrm{s}} \tag{1.1}$$

$$\mathrm{CO_{(g)}} + *_\mathrm{t} \rightleftharpoons \mathrm{CO}^{*_\mathrm{t}} \tag{1.2}$$

$$\mathrm{CO}^{*_\mathrm{s}} + \mathrm{H}^{*_\mathrm{T}} + *_\mathrm{f} + *_\mathrm{t} \rightleftharpoons \mathrm{C\text{-}OH}^{*_\mathrm{f}} + *_\mathrm{s} + *_\mathrm{t} \rightleftharpoons \mathrm{C}^{*_\mathrm{f}} + \mathrm{OH}^{*_\mathrm{t}} + *_\mathrm{s} \tag{1.3}$$

$$\mathrm{CO}^{*_\mathrm{t}} + \mathrm{H}^{*_\mathrm{T}} + *_\mathrm{f} \rightleftharpoons \mathrm{C\text{-}OH}^{*_\mathrm{f}} + *_\mathrm{t} \rightleftharpoons \mathrm{C}^{*_\mathrm{f}} + \mathrm{OH}^{*_\mathrm{t}} \tag{1.4}$$

$$\mathrm{C}^{*_\mathrm{f}} + \mathrm{H}^{*_\mathrm{S}} \rightleftharpoons \mathrm{C\text{-}H}^{*_\mathrm{f}} \rightleftharpoons \mathrm{CH}^{*_\mathrm{f}} \tag{1.5}$$

$$\mathrm{CH^{*f}} + \mathrm{H^{*s}} + *_\mathrm{s} \rightleftharpoons \mathrm{CH\text{-}H^{*s}} + *_\mathrm{f} \rightleftharpoons \mathrm{CH_2^{*s}} + *_\mathrm{f} \tag{1.6}$$

$$\mathrm{CH^{*f}} + \mathrm{H^{*s}} + *_\mathrm{t} \rightleftharpoons \mathrm{CH\text{-}H^{*f}} + *_\mathrm{t} \rightleftharpoons \mathrm{CH_2^{*t}} + *_\mathrm{f} \tag{1.7}$$

$$\mathrm{CH_2^{*s}} + \mathrm{H^{*s}} \rightleftharpoons \mathrm{CH_2\text{-}H^{*s}} \rightleftharpoons \mathrm{CH_3^{*s}} \tag{1.8}$$

$$\mathrm{CH_3^{*s}} + \mathrm{H^{*s}} \rightleftharpoons \mathrm{CH_3\text{-}H^{*s}} \rightleftharpoons \mathrm{CH_{4(g)}} + *_\mathrm{s} \tag{1.9}$$

$$\mathrm{O^{*s}} + \mathrm{H^{*s}} \rightleftharpoons \mathrm{O\text{-}H^{*s}} \rightleftharpoons \mathrm{OH^{*s}} \tag{1.10}$$

$$\mathrm{OH^{*s}} + \mathrm{H^{*s}} \rightleftharpoons \mathrm{H\text{-}OH^{*s}} \rightleftharpoons \mathrm{H_2O_{(g)}} + *_\mathrm{s} \tag{1.11}$$

$$\mathrm{CH^{*f}} + \mathrm{CO^{*s}} \rightleftharpoons \mathrm{CH\text{-}CO^{*s}} + *_\mathrm{f} \rightleftharpoons \mathrm{CHCO^{*s}} + *_\mathrm{f} \tag{1.12}$$

$$\mathrm{CHCO^{*s}} + \mathrm{H^{*s}} \rightleftharpoons \mathrm{H\text{-}CHCO^{*s}} \rightleftharpoons \mathrm{CH_2CO^{*s}} \tag{1.13}$$

$$\mathrm{CH_2CO^{*s}} + \mathrm{H^{*s}} \rightleftharpoons \mathrm{H\text{-}CH_2CO^{*s}} \rightleftharpoons \mathrm{CH_3CO^{*s}} \tag{1.14}$$

$$\mathrm{CH_3CO^{*s}} + \mathrm{H^{*s}} \rightleftharpoons \mathrm{H\text{-}CH_3CO^{*s}} \rightleftharpoons \mathrm{CH_3CHO^{*s}} \tag{1.15}$$

$$\mathrm{CH_3CHO^{*s}} + \mathrm{H^{*s}} \rightleftharpoons \mathrm{CH_3CHO\text{-}H^{*s}} \rightleftharpoons \mathrm{CH_3CHOH^{*s}} \tag{1.16}$$

$$\mathrm{CH_3CHO^{*s}} \rightleftharpoons \mathrm{CH_3CHO_{(g)}} + *_\mathrm{s} \tag{1.17}$$

$$\mathrm{CH_3CHO^{*t}} \rightleftharpoons \mathrm{CH_3CHO_{(g)}} + *_\mathrm{t} \tag{1.18}$$

$$\mathrm{CH_3CHOH^{*s}} + \mathrm{H^{*s}} \rightleftharpoons \mathrm{CH_3CHOH\text{-}H^{*s}} \rightleftharpoons \mathrm{CH_3CH_2OH_{(g)}} + *_\mathrm{s} \tag{1.19}$$

$$\mathrm{CO^{*t}} + \mathrm{H^{*T}} \rightleftharpoons \mathrm{H\text{-}CO^{*t}} \rightleftharpoons \mathrm{CHO^{*t}} \tag{1.20}$$

$$\mathrm{CHO^{*t}} + \mathrm{H^{*T}} \rightleftharpoons \mathrm{HCO\text{-}H^{*t}} \rightleftharpoons \mathrm{CHOH^{*t}} \tag{1.21}$$

$$\mathrm{CHOH^{*t}} + *_\mathrm{t} \rightleftharpoons \mathrm{CH\text{-}OH^{*t}} + *_\mathrm{t} \rightleftharpoons \mathrm{CH^{*t}} + \mathrm{OH^{*t}} \tag{1.22}$$

$$\mathrm{CH^{*t}} + \mathrm{H^{*T}} \rightleftharpoons \mathrm{CH\text{-}H^{*t}} \rightleftharpoons \mathrm{CH_2^{*t}} \tag{1.23}$$

$$\mathrm{CH_2^{*t}} + \mathrm{H^{*T}} \rightleftharpoons \mathrm{H\text{-}CH_2^{*t}} \rightleftharpoons \mathrm{CH_3^{*t}} \tag{1.24}$$

$$\mathrm{CH_3^{*t}} + \mathrm{H^{*T}} \rightleftharpoons \mathrm{H\text{-}CH_3^{*t}} \rightleftharpoons \mathrm{CH_{4(g)}} + *_\mathrm{t} \tag{1.25}$$

$$\mathrm{OH^{*t}} + \mathrm{H^{*T}} \rightleftharpoons \mathrm{H\text{-}OH^{*t}} \rightleftharpoons \mathrm{H_2O_{(g)}} + *_\mathrm{t} \tag{1.26}$$

$$\mathrm{CH^{*t}} + \mathrm{CO^{*t}} \rightleftharpoons \mathrm{CH\text{-}CO^{*t}} + *_\mathrm{t} \rightleftharpoons \mathrm{CHCO^{*t}} + *_\mathrm{t} \tag{1.27}$$

$$\mathrm{CHCO^{*t}} + \mathrm{H^{*T}} \rightleftharpoons \mathrm{H\text{-}CHCO^{*t}} \rightleftharpoons \mathrm{CH_2CO^{*t}} \tag{1.28}$$

$$\mathrm{CH_2CO^{*t}} + \mathrm{H^{*T}} \rightleftharpoons \mathrm{H\text{-}CH_2CO^{*t}} \rightleftharpoons \mathrm{CH_3CO^{*t}} \tag{1.29}$$

$$\mathrm{CH_3CO^{*_t}} + \mathrm{H^{*_T}} \rightleftharpoons \mathrm{H\text{-}CH_3CO^{*_t}} \rightleftharpoons \mathrm{CH_3CHO^{*_t}} \tag{1.30}$$

$$\mathrm{CH_3CHO^{*_t}} + \mathrm{H^{*_T}} \rightleftharpoons \mathrm{CH_3CHO\text{-}H^{*_t}} \rightleftharpoons \mathrm{CH_3CHOH^{*_t}} \tag{1.31}$$

$$\mathrm{CH_3CHOH^{*_t}} + \mathrm{H^{*_T}} \rightleftharpoons \mathrm{CH_3CHOH\text{-}H^{*_t}} \rightleftharpoons \mathrm{CH_3CH_2OH^{*_t}} \tag{1.32}$$

$$\mathrm{CH_3CH_2OH^{*_t}} \rightleftharpoons \mathrm{CH_3CH_2OH_{(g)}} + *_\mathrm{t} \tag{1.33}$$

$$\mathrm{OH^{*_t}} \rightleftharpoons \mathrm{O\text{-}H^{*_t}} \rightleftharpoons \mathrm{O^{*_t}} + \mathrm{H^{*_T}} \tag{1.34}$$

S1.1.2 Diffusion Steps

$$\mathrm{CO^{*_t}} + *_\mathrm{t} \rightleftharpoons \mathrm{CO^{*_t}} + *_\mathrm{t} \tag{1.35}$$

$$\mathrm{CO^{*_t}} + *_\mathrm{s} \rightleftharpoons \mathrm{CO^{*_s}} + *_\mathrm{t} \tag{1.36}$$

$$\mathrm{CO^{*_s}} + *_\mathrm{s} \rightleftharpoons \mathrm{CO^{*_s}} + *_\mathrm{s} \tag{1.37}$$

$$\mathrm{C^{*_f}} + *_\mathrm{f} \rightleftharpoons \mathrm{C^{*_f}} + *_\mathrm{f} \tag{1.38}$$

$$\mathrm{CH^{*_t}} + *_\mathrm{t} \rightleftharpoons \mathrm{CH^{*_t}} + *_\mathrm{t} \tag{1.39}$$

$$\mathrm{CH^{*_t}} + *_\mathrm{f} \rightleftharpoons \mathrm{CH^{*_f}} + *_\mathrm{t} \tag{1.40}$$

$$\mathrm{CH^{*_f}} + *_\mathrm{f} \rightleftharpoons \mathrm{CH^{*_f}} + *_\mathrm{f} \tag{1.41}$$

$$\mathrm{CH_2^{*_t}} + *_\mathrm{t} \rightleftharpoons \mathrm{CH_2^{*_t}} + *_\mathrm{t} \tag{1.42}$$

$$\mathrm{CH_2^{*_t}} + *_\mathrm{s} \rightleftharpoons \mathrm{CH_2^{*_s}} + *_\mathrm{t} \tag{1.43}$$

$$\mathrm{CH_2^{*_s}} + *_\mathrm{s} \rightleftharpoons \mathrm{CH_2^{*_s}} + *_\mathrm{s} \tag{1.44}$$

$$\mathrm{CH_3^{*_t}} + *_\mathrm{t} \rightleftharpoons \mathrm{CH_3^{*_t}} + *_\mathrm{t} \tag{1.45}$$

$$\mathrm{CH_3^{*_t}} + *_\mathrm{s} \rightleftharpoons \mathrm{CH_3^{*_s}} + *_\mathrm{t} \tag{1.46}$$

$$\mathrm{CH_3^{*_s}} + *_\mathrm{s} \rightleftharpoons \mathrm{CH_3^{*_s}} + *_\mathrm{s} \tag{1.47}$$

$$\mathrm{O^{*_t}} + *_\mathrm{t} \rightleftharpoons \mathrm{O^{*_t}} + *_\mathrm{t} \tag{1.48}$$

$$\mathrm{O^{*_t}} + *_\mathrm{s} \rightleftharpoons \mathrm{O^{*_s}} + *_\mathrm{t} \tag{1.49}$$

$$\mathrm{O^{*_s}} + *_\mathrm{s} \rightleftharpoons \mathrm{O^{*_s}} + *_\mathrm{s} \tag{1.50}$$

$$\mathrm{OH}^{*_t} + *_t \rightleftharpoons \mathrm{OH}^{*_t} + *_t \tag{1.51}$$

$$\mathrm{OH}^{*_t} + *_s \rightleftharpoons \mathrm{OH}^{*_s} + *_t \tag{1.52}$$

$$\mathrm{OH}^{*_s} + *_s \rightleftharpoons \mathrm{OH}^{*_s} + *_s \tag{1.53}$$

S1.2 MFA Reaction Network on Rh(211)

The reaction and diffusion steps of the MFA model are identical to the ones of the KMC model except that the hydrogen adsorption to hydrogen reservoir sites with t and s energetics is included

$$\mathrm{H}_{2(g)} + 2*_S \rightleftharpoons 2\mathrm{H}^{*_S} \tag{2.1}$$

$$\mathrm{H}_{2(g)} + 2*_T \rightleftharpoons 2\mathrm{H}^{*_T} \tag{2.2}$$

and same site diffusions are not required. Note that in the KMC model the MFA rate equations for H adsorption/desorption are used to compute an effective H coverage of the T and S sites, which is used instead of treating H as an actual species.

S1.2.1 Original MFA Model by Yang *et al.* on Rh(211)

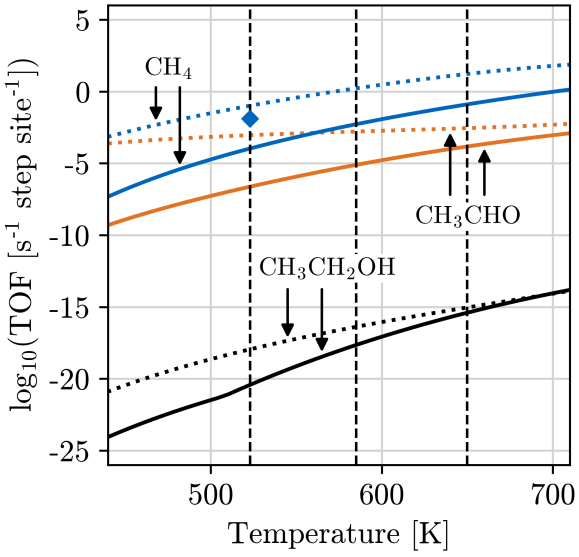


Figure S1: Comparison of the original Rh(211) MFA model by Yang *et al.*[1] with (dotted) and without lateral interactions (solid). The model with interactions is reproduced from their publication and the one without by disabling interactions within the former.

S1.3 KMC Reaction Network on Rh(111)

The KMC Rh(111) model is adopted from Yang *et al.* It is composed of 16 reaction and 5 diffusion steps.

S1.3.1 Reaction Steps

$$\mathrm{CO_{(g)}} + *_\mathrm{t} \rightleftharpoons \mathrm{CO}^{*_\mathrm{t}} \tag{3.1}$$

$$\mathrm{CO}^{*_\mathrm{t}} + \mathrm{H}^{*_\mathrm{T}} \rightleftharpoons \mathrm{H\text{-}CO}^{*_\mathrm{t}} \rightleftharpoons \mathrm{CHO}^{*_\mathrm{t}} \tag{3.2}$$

$$\mathrm{CHO}^{*_\mathrm{t}} + \mathrm{H}^{*_\mathrm{T}} \rightleftharpoons \mathrm{HCO\text{-}H}^{*_\mathrm{t}} \rightleftharpoons \mathrm{CHOH}^{*_\mathrm{t}} \tag{3.3}$$

$$\mathrm{CHOH}^{*_\mathrm{t}} + *_\mathrm{t} \rightleftharpoons \mathrm{CH\text{-}OH}^{*_\mathrm{t}} + *_\mathrm{t} \rightleftharpoons \mathrm{CH}^{*_\mathrm{t}} + \mathrm{OH}^{*_\mathrm{t}} \tag{3.4}$$

$$\mathrm{CH}^{*_\mathrm{t}} + \mathrm{H}^{*_\mathrm{T}} \rightleftharpoons \mathrm{CH\text{-}H}^{*_\mathrm{t}} \rightleftharpoons \mathrm{CH_2}^{*_\mathrm{t}} \tag{3.5}$$

$$\mathrm{CH_2}^{*_\mathrm{t}} + \mathrm{H}^{*_\mathrm{T}} \rightleftharpoons \mathrm{H\text{-}CH_2}^{*_\mathrm{t}} \rightleftharpoons \mathrm{CH_3}^{*_\mathrm{t}} \tag{3.6}$$

$$CH_3^{*t} + H^{*T} \rightleftharpoons H\text{-}CH_3^{*t} \rightleftharpoons CH_{4(g)} + *_t \quad (3.7)$$

$$OH^{*t} + H^{*T} \rightleftharpoons H\text{-}OH^{*t} \rightleftharpoons H_2O_{(g)} + *_t \quad (3.8)$$

$$CH^{*t} + CO^{*t} \rightleftharpoons CH\text{-}CO^{*t} + *_t \rightleftharpoons CHCO^{*t} + *_t \quad (3.9)$$

$$CHCO^{*t} + H^{*T} \rightleftharpoons H\text{-}CHCO^{*t} \rightleftharpoons CH_2CO^{*t} \quad (3.10)$$

$$CH_2CO^{*t} + H^{*T} \rightleftharpoons H\text{-}CH_2CO^{*t} \rightleftharpoons CH_3CO^{*t} \quad (3.11)$$

$$CH_3CO^{*t} + H^{*T} \rightleftharpoons H\text{-}CH_3CO^{*t} \rightleftharpoons CH_3CHO^{*t} \quad (3.12)$$

$$CH_3CHO^{*t} \rightleftharpoons CH_3CHO_{(g)} + *_t \quad (3.13)$$

$$CH_3CHO^{*t} + H^{*T} \rightleftharpoons CH_3CHO\text{-}H^{*t} \rightleftharpoons CH_3CHOH^{*t} \quad (3.14)$$

$$CH_3CHOH^{*t} + H^{*T} \rightleftharpoons CH_3CHOH\text{-}H^{*t} \rightleftharpoons CH_3CH_2OH^{*t} \quad (3.15)$$

$$CH_3CH_2OH^{*t} \rightleftharpoons CH_3CH_2OH_{(g)} + *_t \quad (3.16)$$

S1.3.2 Diffusion Steps

$$CO^{*t} + *_t \rightleftharpoons CO^{*t} + *_t \quad (3.17)$$

$$CH^{*t} + *_t \rightleftharpoons CH^{*t} + *_t \quad (3.18)$$

$$CH_2^{*t} + *_t \rightleftharpoons CH_2^{*t} + *_t \quad (3.19)$$

$$CH_3^{*t} + *_t \rightleftharpoons CH_3^{*t} + *_t \quad (3.20)$$

$$OH^{*t} + *_t \rightleftharpoons OH^{*t} + *_t \quad (3.21)$$

S1.4 MFA Reaction Network on Rh(111)

Like in the Rh(211) model the reaction steps of the MFA model are identical to the KMC model and hydrogen adsorption to hydrogen reservoir sites with t energetics is included

$$H_{2(g)} + 2*_T \rightleftharpoons 2H^{*T} \quad (4.1)$$

This model is identical to the one by Yang *et al.*

S2 Additional DFT Computational Details

DFT calculations were performed as in Yang *et al.*[1] using Quantum Espresso,[6] ultrasoft pseudopotentials, and the BEEF-vdW exchange-correlation functional[7] with a plane wave cutoff of 500 eV and a density cutoff of 5000 eV. The Brillouin zone was sampled with a 4×4×1 Monkhorst-Pack k-point grid. The Rh(211) surface slab was modelled using a (2×1) supercell with nine layers and the bottom six layers fixed in their bulk truncated positions and the Rh(111) surface slabs were modelled using a (3×3) supercell with four layers and the bottom two layers fixed in their bulk truncated positions. The slabs were relaxed until reaching a maximum force threshold of 0.05 eV/Å. A vacuum region of at least 15 Å perpendicular to the surface was used and a dipole correction was applied to all surfaces.[8] Transition state geometries were calculated with the CI-NEB method[9] analogous to Yang *et al.*[1] Vibrational frequencies of adsorbed species below 56 cm^{-1} are set to 56 cm^{-1} to approximate the entropic contribution from pseudotranslational/rotational degrees of freedom.

Table S1: Additional DFT calculations for the modified Rh(211) reaction model.

Species	Surface	Formation energy [eV]	Vibrational frequencies [cm^{-1}]
O^{*t}	Rh(111)	−0.19	359.5, 393.3, 507.0
O-H^{*t}	Rh(111)	1.13	130.2, 179.5, 293.7, 497.8, 975.7
CH-H^{*f}	Rh(211)	0.36	247.4, 286.7, 460.4, 579.8, 794.5, 958.4, 1325.6, 2936.3

Table S2: Nearest neighbor interaction energies (ε) in eV.

Interaction	ε [eV]	Interaction	ε [eV]
$C^{*f}-CO^{*s}$	0.46	$CH_3CO^{*s}-CO^{*s}$	0.07
$CH^{*f}-CO^{*t}$	0.39	$OH^{*t}-CO^{*t}$	0.06
$CH_2CO^{*s}-CO^{*s}$	0.35	$CH_2CO^{*s}-CO^{*t}$	0.06
$C^{*f}-CH^{*f}$	0.31	$CH_3^{*t}-CO^{*s}$	0.05
$CO^{*s}-O^{*s}$	0.29	$CHO^{*t}-CO^{*t}$	0.05
$CH_3CHOH^{*s}-CO^{*s}$	0.27	$CHCO^{*s}-CO^{*t}$	0.05
$C^{*f}-C^{*f}$	0.25	$CH_3^{*s}-CO^{*t}$	0.04
$CH^{*f}-CH^{*f}$	0.25	$CH^{*f}-CO^{*s}$	0.04
$CO^{*s}-OH^{*s}$	0.24	$CHCO^{*s}-CO^{*s}$	0.03
$CH_3CO^{*s}-CO^{*t}$	0.20	$OH^{*t}-CO^{*s}$	0.03
$CO^{*t}-CO^{*t}$	0.19	$CH^{*t}-CO^{*s}$	0.02
$CH_2CO^{*t}-CO^{*t}$	0.15	$CH_2^{*t}-CO^{*s}$	0.01
$CH_3CHOH^{*t}-CO^{*t}$	0.12	$CH_3CO^{*t}-CO^{*t}$	0.01
$CH_3^{*t}-CO^{*t}$	0.10	$CO^{*t}-CO^{*s}$	−0.01
$CO^{*s}-CO^{*s}$	0.10	$CH_3CHO^{*s}-CO^{*t}$	−0.03
$O^{*t}-CO^{*t}$	0.10	$CO^{*t}-O^{*s}$	−0.05
$O^{*t}-CO^{*s}$	0.09	$CH_3CHOH^{*s}-CO^{*t}$	−0.07
$CHOH^{*t}-CO^{*t}$	0.09	$CH_3CHO^{*t}-CO^{*t}$	−0.08
$CH_3CHO^{*s}-CO^{*s}$	0.09	$CH_3CH_2OH^{*t}-CO^{*t}$	−0.09
$CHCO^{*t}-CO^{*t}$	0.09	$CO^{*t}-OH^{*s}$	−0.11
$CH_3^{*s}-CO^{*s}$	0.08	$C^{*f}-CO^{*t}$	−0.17
$CH^{*t}-CO^{*t}$	0.07		

Table S3: Brønsted-Evans-Polanyi scaling parameter (α) values.

α	forward	reverse
$\alpha_{\text{adsorption}}$	0.000	1.000
$\alpha_{\text{diffusion}}$	0.500	0.500
$\alpha_{CO+H\rightarrow C+OH}$	0.792	0.208
$\alpha_{CHOH\rightarrow CH+OH}$	0.792	0.208
$\alpha_{C+H\rightarrow CH}$	0.363	0.637
$\alpha_{CH+H\rightarrow CH_2}$	0.239	0.761
$\alpha_{CH_2+H\rightarrow CH_3}$	0.243	0.757
$\alpha_{CH_3+H\rightarrow CH_4}$	0.211	0.789
$\alpha_{\text{C hydrogenation}}$	0.264	0.736
$\alpha_{O+H\rightarrow OH}$	0.563	0.437
$\alpha_{OH+H\rightarrow H_2O}$	0.350	0.650
$\alpha_{\text{O hydrogenation}}$	0.456	0.544
$\alpha_{CO+CH\rightarrow CHCO}$	0.204	0.796
$\alpha_{H+CO\rightarrow CHO}$	0.562	0.438

S3 Additional KMC Computational Details

All KMC simulations were performed with `kmos`[10] using the on-the-fly (otf) backend[11] and the acceleration algorithm described by Dybeck *et al.* as implemented in `kmos`.[5,12,13] Starting with a `random_seed` of 1 for all simulations the systems were relaxed for $5 \cdot 10^7$ steps and subsequently checked every $5 \cdot 10^7$ steps until either the runtime exceeded one week of simulation or 11 molecules of $CH_3CHO_{(g)}$ for the Rh(111) facet (26 molecules for the Rh(211) facet) were formed. In the acceleration algorithm there are several parameters to set. Two of the parameters, n_e and δ, were fixed to the same values as used in our previous work.[5] That is, reaction channels need to be executed $n_e = 200$ times to be considered locally quasi-equilibrated and the threshold parameter is set to $\delta = 0.2$. Furthermore, to yield converged results different buffer parameters N_f and lattice sizes have been tested as shown in Figure S2. Based on these results, we chose a 10×10 lattice with $N_f = 25$ for Rh(211) and a 25×25 lattice with $N_f = 50$ for Rh(111).

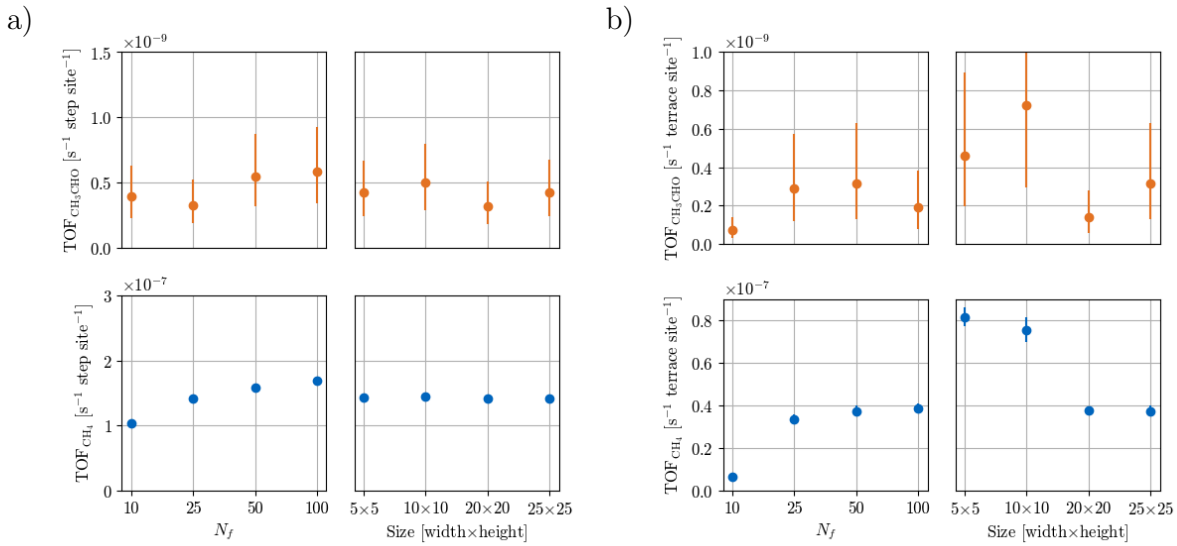


Figure S2: Acceleration parameter analysis for the formation of CH_4 and CH_3CHO on (a) Rh(211) and (b) Rh(111) with lateral interactions at 523 K. Left plots: Different buffer parameters N_f for (a) a 30×30 Rh(211) lattice and (b) a 25×25 Rh(111) lattice. Right plots: Different lattice sizes for (a) $N_f = 25$ and (b) $N_f = 50$.

S3.1 Sensitivities of the Rh(211) KMC Model with Lateral Interactions

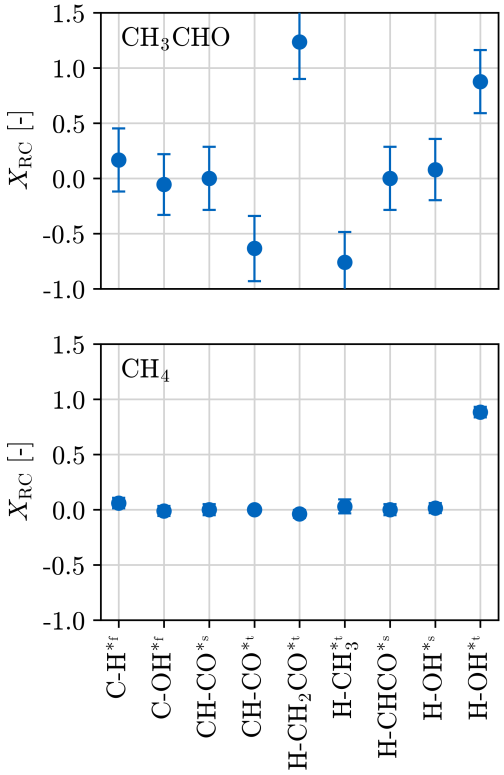


Figure S3: Sensitivity analysis of the KMC model for Rh(211) with lateral interactions, showing the DRC for $CH_{4,(g)}$ (bottom) and $CH_3CHO_{(g)}$ (top) of selected TSs. The analysis is performed at 650 K

S3.2 Lateral Interaction Contributions Within the CO Activation Step

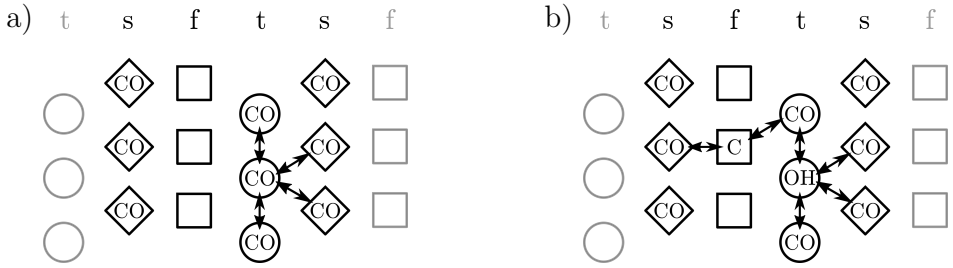


Figure S4: Most probable configurations for the IS (a) and FS (b) of the CO activation reaction on the Rh(211) facet.

S3.3 Effect of Lateral Interactions on Coverage within KMC

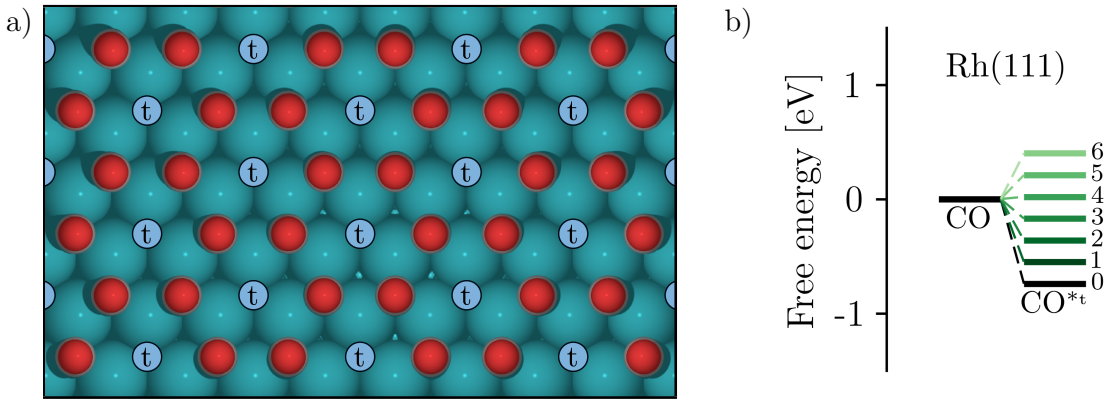


Figure S5: (a) Coverage pattern of CO in the Rh(111) KMC model with lateral interactions. Due to the interactions a honeycomb-like adsorption pattern is formed with 2/3 coverage of CO and 1/3 empty t sites. (b) Free energy levels of the CO^{*t} state with 0–6 neighboring sites occupied by other CO molecules. The analysis is performed at 650 K.

Table S4: Coverages of the different Rh(211) and Rh(111) KMC models.

Model	Temperature	Site	Coverage [%] CO	C	empty	rest
Rh(211) w/ lateral interactions	523	t	100.0	–	0.0	0.0
		s	100.0	–	0.0	0.0
		f	–	5.1	94.9	0.0
	585	t	100.0	–	0.0	0.0
		s	100.0	–	0.0	0.0
		f	–	0.1	99.9	0.0
	650	t	99.9	–	0.1	0.0
		s	100.0	–	0.0	0.0
		f	–	0.0	100.0	0.0
Rh(211) w/o lateral interactions	523	t	100.0	–	0.0	0.0
		s	100.0	–	0.0	0.0
		f	–	0.8	99.0	0.2
	585	t	100.0	–	0.0	0.0
		s	100.0	–	0.0	0.0
		f	–	0.7	99.2	0.1
	650	t	100.0	–	0.0	0.0
		s	100.0	–	0.0	0.0
		f	–	0.6	99.3	0.1
Rh(111) w/ lateral interactions	523	t	67.8	–	32.2	0.0
	585	t	65.9	–	34.1	0.0
	650	t	62.0	–	38.0	0.0

S3.4 Sensitivities of the KMC and MFA Models for Rh(111) with Lateral Interactions

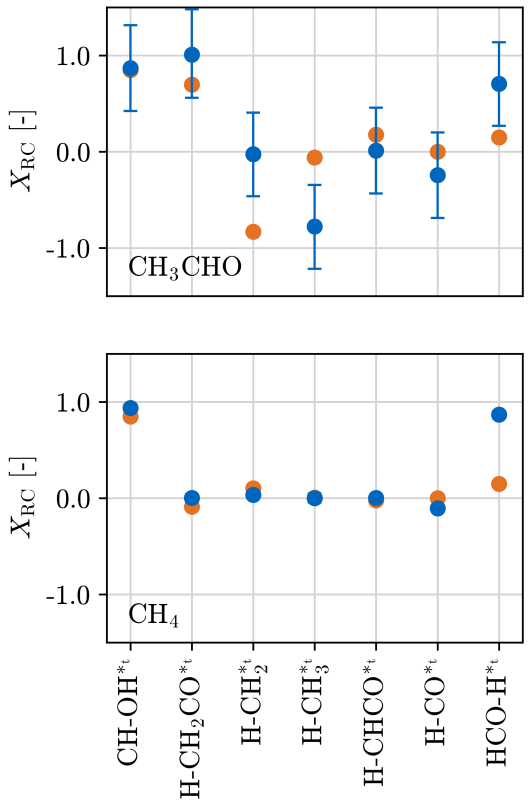


Figure S6: Sensitivity analysis of the KMC (blue points) and MFA (orange points) models for Rh(111) with lateral interactions, showing the DRC for $CH_{4,(g)}$ (bottom) and $CH_3CHO_{(g)}$ (top) of selected TSs. The analysis is performed at 650 K.

References

(1) Yang, N.; Medford, A. J.; Liu, X.; Studt, F.; Bligaard, T.; Bent, S. F.; Nørskov, J. K. Intrinsic Selectivity and Structure Sensitivity of Rhodium Catalysts for C_{2+} Oxygenate Production. *J. Am. Chem. Soc.* **2016**, *138*, 3705–3714.

(2) 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.

(3) Ferrin, P.; Simonetti, D.; Kandoi, S.; Kunkes, E.; Dumesic, J. A.; Nørskov, J. K.; Mavrikakis, M. Modeling Ethanol Decomposition on Transition Metals: A Combined Application of Scaling and Brønsted-Evans-Polanyi Relations. *J. Am. Chem. Soc.* **2009**, *131*, 5809–5815.

(4) Choi, Y.; Liu, P. Mechanism of Ethanol Synthesis from Syngas on Rh(111). *J. Am. Chem. Soc.* **2009**, *131*, 13054–13061.

(5) 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.

(6) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Corso, A. D.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.

(7) 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* **2012**, *85*, 235149.

(8) Bengtsson, L. Dipole correction for surface supercell calculations. *Phys. Rev. B* **1999**, *59*, 12301–12304.

(9) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A climbing image nudged elastic band

method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113*, 9901–9904.

(10) Hoffmann, M. J.; Matera, S.; Reuter, K. kmos: A lattice kinetic Monte Carlo framework. *Comput. Phys. Commun.* **2014**, *185*, 2138–2150.

(11) Lorenzi, J. M. Ab initio modeling of heterogeneous catalysis: Multiscale analysis of interacting reaction pathways. Ph.D. thesis, Technical University of Munich, 2017.

(12) Dybeck, E. C.; Plaisance, C. P.; Neurock, M. Generalized Temporal Acceleration Scheme for Kinetic Monte Carlo Simulations of Surface Catalytic Processes by Scaling the Rates of Fast Reactions. *J. Chem. Theory Comput.* **2017**, *13*, 1525–1538.

(13) Seibt, M. Temporal Acceleration of Kinetic Monte Carlo Simulations Applied to Heterogeneous Catalysis. M.Sc. thesis, Technical University of Munich, 2018.