Adsorbate Partition Functions via Phase Space Integration: Quantifying the Effect of Translational Anharmonicity on Thermodynamic Properties

Katrín Blöndal,[†] Khachik Sargsyan,[‡] David H. Bross,[¶] Branko Ruscic,[¶] and C. Franklin Goldsmith*,[†]

[†]*Chemical Engineering Group, School of Engineering, Brown University, Providence, RI 02912, USA*

[‡]*Sandia National Laboratories, Livermore, CA 94550, USA*

[¶]*Chemical Sciences and Engineering Division, Argonne National Laboratory, Lemont, IL 60439, USA*

E-mail: franklin_goldsmith@brown.edu

Abstract

A new method for computing anharmonic thermophysical properties for adsorbates on metal surfaces is presented. Classical Monte Carlo phase space integration is performed to calculate the partition function for the motion of a hydrogen atom on Cu(111). A minima-preserving neural network potential energy surface is used within the integration routine. Two different sampling schema for generating the training data are presented, and two different density functionals are used. The results are benchmarked against direct state counting results using discrete variable representation. The phase space integration results are in excellent quantitative agreement with the benchmark results. Additionally, both the discrete variable representation and the phase space integration results confirm that the motion of H on Cu(111) is highly anharmonic. The results were applied to calculate the free energy of dissociative adsorption of H_2 and the resulting Langmuir isotherms at 400, 800 and 1200 K in a partial pressure range of 0-1 bar. It shows that the anharmonic effects lead to significantly higher predicted surface site fractions of hydrogen.

Introduction

Microkinetic modeling is an increasingly popular and effective way to simulate catalytic processes and optimize catalyst design and operating conditions.[1] At a minimum, a microkinetic mechanism must be thermodynamically consistent.[2] Thermodynamic consistency requires reaction equilibrium constants, which in turn requires accurate free energies of adsorbates. In all but the rarest cases, adsorbate free energies cannot be obtained easily experimentally. Instead, theoretical methods are utilized to obtain partition functions of the adsorbates, from which their thermophysical properties, such as the enthalpy increment, entropy, and heat capacity, can be derived.

Arguably the most common model for adsorbate partition functions is the harmonic oscillator (HO) model, also known as the 2D lattice model.[3] The harmonic oscillator model

is particularly suitable for tightly bound species. However, the HO approximation is in poor agreement with experimental data; for example, it tends to underestimate adsorbate entropy.[4] For weakly bound (*e.g.* physisorbed) species, which have low diffusion barriers, an alternative model is the free translator (FT), also known as the 2D gas model.[5] Collectively, the harmonic oscillator and the free translator models represent idealized limits (infinite barriers and no barriers, respectively) and do not accurately describe most systems at finite temperatures. Recent studies[6,7] show that anharmonic motion (which, in this case, arises from finite barriers) is important for accurate entropies and thence free energies in heterogeneous catalysis.

The recently developed hindered translator (HT) model[8] bridges the gap between the HO and the FT models with an analytical expression for the potential energy surface (PES) that is sinusoidal in two degrees of freedom parallel to the surface. It provides a more accurate result while still being comparatively easy to implement. An alternative way to bridge the gap is with complete potential energy sampling (CPES),[9] introduced by Jørgensen and Grönbeck in 2017, which has shown to perform similarly to the hindered translator model. CPES treats in-plane adsorbate motion as a classical continuous system and is more general than the hindered translator model, as it takes into account the detailed energy landscape. Jørgensen and Grönbeck later presented Monte Carlo complete potential energy sampling (MCPES),[10] a computational method for calculating molecular entropy in zeolites. In MCPES, the partition function includes anharmonic contributions to all transitional degrees of freedom of the molecule in the zeolite and is obtained semiclassically, by integration of the potential energy surface.

The HO, FT, HT and CPES free energy approximations for adsorbates on flat surfaces were benchmarked against more accurate free energies directly computed from first-principles for monatomic adsorbates in a study by Bajpai et al.[11] It was found that the commonly used analytical models, HO, FT, and HT, often do not properly capture adsorbate free energies, and that their performance depends on the shape and separability of the PES. Moreover,

the Bajpai et al. study found that a classical treatment for in-plane motion such as CPES can be used to evaluate the free energies of monatomic adsorbates accurately, except at low temperatures.

Potential energy sampling based methods are therefore promising for accurate free energies of arbitrary adsorbates at high temperatures. In this study we use a sampling based approach to obtain the coupled, anharmonic partition function of a hydrogen atom on a Cu(111) surface. Specifically, the three translational degrees of freedom are all treated as coupled and anharmonic. We use a classical phase space integral representation for the anharmonic partition function and solve it with Monte Carlo (MC) sampling.

The MC sampling is exact (in the classical limit), given a sufficient number of sampling points. In principle, it is possible to perform the direct sampling using density functional theory (DFT); in practice, however, direct DFT sampling is impractical. A suitably dense sampling may require hundreds of thousands of discrete energy evaluations, which could be computationally prohibitive using DFT. Instead, we develop an accurate surrogate potential energy surface using machine learning methods. This machine learning surrogate PES can be evaluated quickly during the integration process with negligible loss in accuracy, compared to DFT.

A handcrafted form inspired by the quadratic approximation is used to construct the potential energy surrogate. This functional form, the minima-preserving neural network (MP-NN), requires the Hessian matrix and the optimized position of the adsorbate at each minimum. The only features that are required in this machine learning method are the distances of the hydrogen atom from each minimum. This machine learning method does not require elemental information, which distinguishes it from the more commonly used fingerprinting methods. For example, the most frequently used methods involve either using handcrafted mathematical expressions to describe local environments of atoms, such as Behler-Parrinello (BP) symmetry functions,[12] or using convolutional layers involving distances and atom features, such as lattice convolutional neural networks (LCNN).[13] Appli-

cations of machine learning for atomistic potential energy surfaces is a rapidly developing field, and a thorough review is outside the scope of the present work. Please refer to review studies such as Refs. 14, 15 and 16.

We consider the performance of the MP-NN surrogate potential energy surface in phase space integration (PSI) in evaluating the partition function and compare our result to the common approximations, HO, FT, and HT. We quantify the effect of anharmonicity, looking at the anharmonic correction factor f, analogous to how anharmonicity effects have been quantified for vibrational modes of gas-phase species.[17] Furthermore, we directly derive the entropy, enthalpy increment, and heat capacity from the partition function. Lastly, we calculate the standard-state Gibbs free energy and plot the resulting Langmuir isotherms for H_2(g) dissociative adsorption, and compare to the standard models.

Methods

Electronic Structure Theory

Density Functional Theory (DFT) calculations were performed with the plane-wave code Quantum ESPRESSO.[18–20] Two different functionals were used in conjunction with two different methods for generating training data (described in greater detail in the next section). In the first set of calculations, the Perdew-Burke-Ernzerhof functional[21] was used with Grimme's D3 corrections[22] including three-body Axilrod-Teller-Muto corrections,[23,24] PBE-D3(ABC). PAW pseudopotentials generated using the "atomic" code by Dal Corso[25,26] were used.

The optimized lattice constant with the PBE-D3(ABC) functional was found to be 3.58 Å. The H-Cu(111) system was modeled using a 4-layer Cu(111) slab with a 3 × 3 unit cell and a 17 Å vacuum between the H atop and the bottom of the repeated image. All the Cu surface atoms were held fixed in their bulk geometry while the H atom was relaxed on the surface using the LBFGS algorithm in the Atomic Simulation Environment (ASE),[27]

with a maximum force convergence criterion of 0.01 eV/Å along all directions. Energies were calculated with a plane-wave cutoff of 60 Ry and a (6×6×1) k-point grid, criteria at which the H adsorption energy was converged to within 0.09 eV. The Marzari-Vanderbilt smearing method,[28] with a smearing width of 0.005 Ry, was used. Spin polarization was not considered in the calculations. The fcc binding site was found to be the optimal position of H on the Cu(111) surface. A finite-difference approximation was used to obtain the Hessian matrix of the H atom in the fcc and the hcp binding sites.

The potential energy $V(x,y,z)$ was calculated at total of 14,395 (x,y,z) positions of H within the area of the rhombus-shaped primitive unit cell, shown in Figure 1. The z values range from 0.5 Å below the optimal z value of H in the fcc position up to 3.5 Å above it. The training data includes the four stationary points of H on the surface, the fcc, hcp, bridge and atop sites. A small portion of the data, or 1,024 geometries, were distributed over the whole constrained volume as a Sobol sequence.[29] The remaining 13,367 geometries were sampled from a multivariate normal (MVN) distribution with a covariance set to be equal to the inverse Hessian matrix of the H atom in the fcc position. This procedure was to ensure ample sampling close to the minimum energy, the highest Boltzmann weight area (see the Supplemental Information for details).

In the second set of calculations, new training data were computed with the BEEF-vdW functional[30–32] using the LIBBEEF package,[33] which is automatically enabled in Quantum ESPRESSO versions 6.6 and higher. We used SG15 Optimized Norm-Conserving Vanderbilt (ONCV) pseudopotentials that were generated using the "ONCVPSP" code by Hamann.[34–36] The optimized lattice constant of Cu was found to be 3.66 Å. The same procedure as for the PBE-D3(ABC) functional was mostly followed with a few modifications. A smearing width of 0.02 Ry was used and the H adsorption energy was converged to within 0.03 eV using the 60 Ry and a (6×6×1) k-point grid.

The BEEF-vdW training data consisted of 9,382 geometries within the primitive unit cell area, where the z values range from 0.5 Å below the optimal z value of H in the fcc

position up to 4.25 Å above it. This set contains the fcc, hcp, bridge and atop sites as well as 1,024 samples according to a Sobol sequence. Unlike the PBE-D3(ABC) set, the remaining samples belong to two MVN distributions: 4,122 of the samples belong to an MVN centered at the fcc position with the associated inverse-Hessian covariance, and 4,232 samples belong to another MVN distribution centered at the hcp position with its corresponding inverse-Hessian as a covariance matrix. Using two MVN sample sets is an attempt at more efficient sampling for maximum PES surrogate accuracy close to the minimum energy positions. This dual-minima approach is expected to be an improvement, especially in this case where the hcp position was found to be effectively equivalent to the fcc position, with a less than 1 meV difference in potential energy between the two minima.

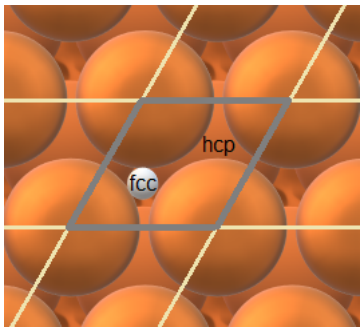

Figure 1: H in the fcc binding site of a Cu(111) surface and the periodic primitive unit cell area that contains all training positions.

Surrogate Potential Energy Surface

As described in the introduction, the present approach uses a surrogate potential energy surface: $V_s(x, y, z) \approx V(x, y, z)$, which is constructed using the training set of $\boldsymbol{x} = (x, y, z)$ geometries and corresponding DFT energies described in the previous section. In other words, the training set is $\{(\boldsymbol{x}^{(j)}, V(\boldsymbol{x}^{(j)}))\}$ for $j = 1, \ldots, N$, where N=14,394 for the PBE-D3(ABC) case and N=9,380 for the BEEF-vdW case. We removed the training sample corresponding to the fcc minimum from the PBE-D3(ABC) set, and two samples corresponding to the fcc and hcp minima from the BEEF-vdW set. This was done to ensure that

the training sample transformation, described later in this section, is well-defined, avoiding division by zero. As for the parametric form of the surrogate $V_s(x, y, z)$, we employ neural networks (NNs) within PyTorch machine learning framework, with special adjustments to ensure accuracy of the surrogate near the fcc and the hcp binding sites of hydrogen on the surface. Because the Boltzmann weighting will be dominated by the lowest energy configurations, we wish to ensure that the surrogate is most accurate in the proximity of the various minima. Accordingly, we introduce the minima-preserving neural network (MP-NN) surrogate. The MP-NN surrogate is inspired by the quadratic approximation near a minimum $\boldsymbol{x}_0 = (x_0, y_0, z_0)$. Given the minimum $V(\boldsymbol{x}_0)$ and the Hessian $H(\boldsymbol{x}_0)$, consider a surrogate of a form

$$V_s(\boldsymbol{x}) = V(\boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0)^T H(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)e^{g(\boldsymbol{x}-\boldsymbol{x}_0)}. \tag{1}$$

By construction, $V_s(\boldsymbol{x})$ has a minimum $V(\boldsymbol{x}_0)$ at $\boldsymbol{x}_0$, and has a positive definite Hessian. In order to restrict our surrogate to the form given in Eq. (1), we strive to learn the general function $g(\boldsymbol{x})$, and transform the training samples according to

$$\begin{aligned} \tilde{\boldsymbol{x}}^{(j)} &= \boldsymbol{x}^{(j)} - \boldsymbol{x}_0 \\ g(\tilde{\boldsymbol{x}}^{(j)}) &= \log\left\{ \frac{V(\boldsymbol{x}^{(j)}) - V(\boldsymbol{x}_0)}{\frac{1}{2}(\boldsymbol{x}^{(j)} - \boldsymbol{x}_0)^T H(\boldsymbol{x}_0)(\boldsymbol{x}^{(j)} - \boldsymbol{x}_0)} \right\} \end{aligned} \tag{2}$$

and, using the transformed training set $\{(\tilde{\boldsymbol{x}}^{(j)}, g(\tilde{\boldsymbol{x}}^{(j)}))\}$ for $j = 1, \ldots, N$, we learn the multiplicative correction factor exponent as a neural network $g(\boldsymbol{x}) = \mathrm{NN}(\boldsymbol{x})$. In the case of two minima, $\boldsymbol{x}_0^{(1)}$ and $\boldsymbol{x}_0^{(2)}$, we construct surrogates of form (1) for each of them individually using only training evaluations within a certain distance from the corresponding minimum, arriving at $V_s^{(1)}(\boldsymbol{x})$ and $V_s^{(2)}(\boldsymbol{x})$, each being accurate near the appropriate minima. Then we obtain the final MP-NN form of the surrogate as a linear combination of these two surrogates:

$$
\begin{aligned}
w_i(\boldsymbol{x}) &= \frac{e^{-d(\boldsymbol{x},\boldsymbol{x}_0^{(i)})/\epsilon}}{e^{-d(\boldsymbol{x},\boldsymbol{x}_0^{(1)})/\epsilon} + e^{-d(\boldsymbol{x},\boldsymbol{x}_0^{(2)})/\epsilon}} \\
V_s^{\text{joint}}(\boldsymbol{x}) &= w_1(\boldsymbol{x}) V_s^{(1)}(\boldsymbol{x}) + w_2(\boldsymbol{x}) V_s^{(2)}(\boldsymbol{x})
\end{aligned}
\tag{3}
$$

where $d(\boldsymbol{x}, \boldsymbol{x}_0^{(i)})$ within the weight $w_i(\boldsymbol{x})$ is the Euclidean distance of the point $\boldsymbol{x}$ from the i-th minimum ($i = 1, 2$), and ϵ is a small positive 'smoothing' factor. This distance-weighted form ensures the transition from the region near one minimum to a region near another minimum, with the corresponding surrogates switching their importance in the linear combination. The cost is a slight deterioration of minimum constraints, but small ϵ and large distance between minima reduces this effect. The sketch of the MP-NN surrogate form is illustrated in Figure 2. The overall surrogate is a weighted linear combination (WLC) of two separately trained surrogates. Each of the two identical paths in Figure 2 represents a surrogate corresponding to a minimum, indicating operations of shifting input, training a NN and exponentiating it as a correction to a quadratic, before a vertical shift. The operations are equivalent to transforming the training input/output pairs according to Eq. (2).

Two different training schema are considered. In the first approach, "surrogate 1", which is used for the PBE-D3(ABC) functional, the NN architecture that plays the role of the multiplicative correction (see Figure 2) was set to have 3 hidden layers with 111 units for each, using a Rectified Linear Unit (ReLU) activation function. The NN is trained via Adam optimization,[37] using 5,000 epochs with learning rate 0.01 and a batch size of 200. For the second training scheme, "surrogate 2", which is used for the BEEF-vdW training data, the NN architecture was set to have 3 hidden layers with 1,111 units for each, using a hyperbolic tangent activation function. Similar to the previous case, the NN is trained via Adam optimization, using 5,000 epochs with learning rate 0.01 and a batch size of 200. In other words, the second training approach has 10x more units. The other salient difference between the two approaches is with respect to the generation of the training data

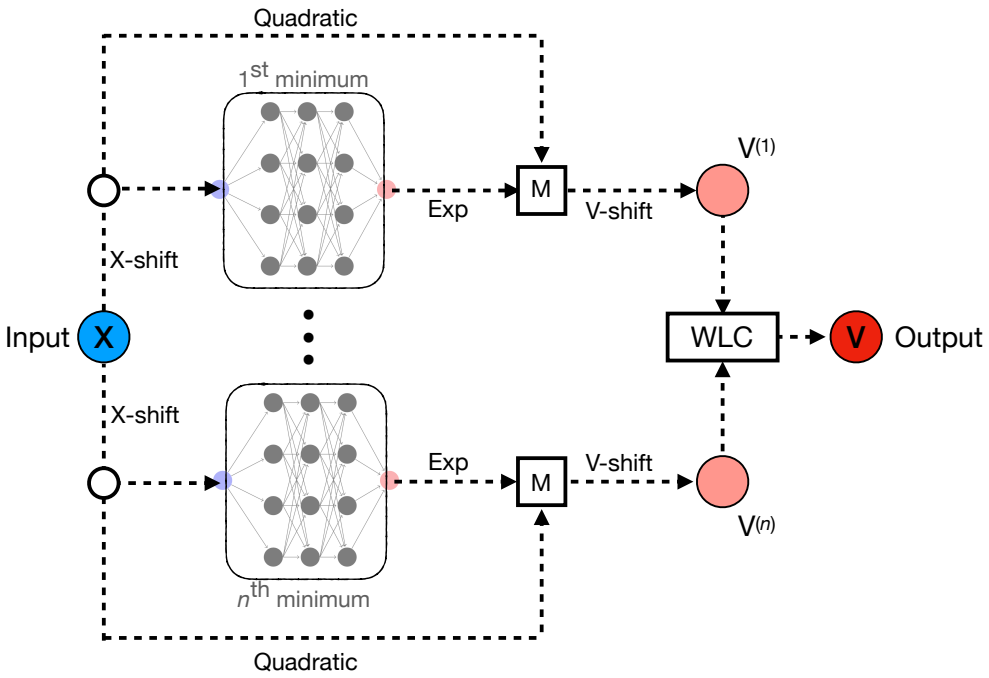

Figure 2: The overall architecture of the MP-NN surrogate, as a weighted linear combination (WLC) of minima-preserving surrogates in the form of NN-based multiplicative corrections to a quadratic function. M indicates a multiplication, while X-shift and V-shift indicate input and output shifts, respectively. In the present work, the two minima for which neural networks are created are the fcc and hcp.

itself. For the PBE-D3(ABC) functional, only the fcc minimum was considered for the MVN distribution. For the BEEF-vdW functional, both the fcc and the hcp minima were explicitly considered. As will be demonstrated in the Results section, including the neighborhoods of both minima in the training scheme and increasing the number of units improves the performance, relative to the first case. Note that the MP-NN surrogate can be reformulated and applied to any coordinates or features. While here we demonstrate it for the Cartesian coordinates, it can easily be adapted to work with transformed coordinates or features, such as, *e.g.*, Behler-Parrinello symmetry functions, as long as the minima locations and the corresponding Hessians are transformed to be rewritten with respect to the new coordinates.

Direct State Counting

Direct state counting of the three degrees of freedom was used to obtain a reference set of thermodynamic functions. Explicit eigenvalues were obtained for the three degrees of freedom (x, y, z) of the adsorbate relative to the surface using a discrete variable representation (DVR) method.[38] A direct symmetric eigensolver from LAPACK was implemented in a local version of NITROGEN[39] package for the dense eigensolve. The direct count of eigenstates from 2D scans of x, y translation along the surface, as well as 1D translation away from the surface in z, were solved via DVR basis functions that corresponded to Cartesian basis functions in these coordinates. The x, y-surface was calculated on a rectangular grid of 93×81 grid points spanning 2.531 angstroms in x and 2.1919 angstroms in y using the PBE-D3 (ABC) MP-NN surrogate. The BEEF-vdW based surrogate used 93×81 grid points that spanned 2.588 angstroms in x and 2.2412 angstroms in y. These x, y grids had periodic boundary conditions enforced through a Fourier DVR basis functions that maintained translational invariance at the endpoints. The z coordinate was relaxed on the x, y surface. The 1D DVR of the z coordinate was scanned from 0.1 angstrom to 2.0 angstroms above the surface with 1,000 grid points on both surrogates. Convergence tests were carried out on grids with up to double the number of grid points. The changes to the density of states from larger grids were minimal and the resulting partition functions at the highest calculated temperature differed by less than 1.0×10^{-5}. The product of the resulting x, y and z states were used in thermodynamic summation to compute the quantum mechanical thermodynamic functions, see Equations (15–17) below for details.

Monte Carlo Phase Space Integration

The statistical analysis begins with the classical Hamiltonian for 3D translation of the adsorbate species relative to the catalyst surface:

$$H(p_x, p_y, p_z, x, y, z) = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} + V(x, y, z) \tag{4}$$

where p_i is the momentum in the i^{th} direction, m is the mass, and (x,y,z) are the positions of the center of mass of the H atom on the Cu(111) surface. The corresponding classical partition function for this system is the Boltzmann factor of the Hamiltonian, integrated over the momentum and position. The momentum integral is a Gaussian integral, which can be solved separately from the potential part. Due to the periodicity of the surface, the limits of integration are restricted to the area of the unit cell and the finite distance above the surface. The resulting classical partition function is:

$$q^{\text{classical}} = \frac{(2\pi\, m\, k_B T)^{3/2}}{h^3} \int_{z_1}^{z_2} \int_{y_1}^{y_2} \int_{x_1}^{x_2} e^{-\beta\, V(x,y,z)}\, dx\, dy\, dz, \tag{5}$$

where $\beta = 1/(k_B T)$ and k_B is the Boltzmann constant, x_1, x_2, y_1, y_2 are the integration limits of the repeatable unit cell area, and z_1 and z_2 are the limits in the direction perpendicular to the surface. The potential energy in Equation (5) is not approximated with an analytical expression but rather is computed directly from the surrogate potential; consequently, the integral is evaluated numerically via Monte Carlo integration:

$$q^{\text{classical}} \approx \frac{(2\pi\, m\, k_B T)^{3/2}}{h^3} \frac{\Omega}{N} \sum_{i=1}^{N} e^{-\beta V_s(x,y,z)} \tag{6}$$

Here, Ω represents the total volume of the sampled region, and N is the number of sampled configurations. The Monte Carlo PSI (Equation 6) was performed on the MP-NN surrogate with the integration convergence parameter

$$c \equiv \left| \frac{q_{\text{old}}^{\text{classical}} - q_{\text{new}}^{\text{classical}}}{q_{\text{old}}^{\text{classical}}} \right| \tag{7}$$

which calculates the change in $q^{\text{classical}}$ upon adding 5 new geometries to the set where q_{old} is the partition function without the five most recently added geometries (that are included in q_{new}). This parameter is checked every 5 new geometries for each temperature T, which in this study encompasses 31 T values in the range 290-1400 K. The convergence criterion was set to $c < 1.0\times10^{-6}$.

Thermophysical Properties

For the present work, the three most relevant thermophysical properties that are obtained from the partition function are the enthalpy increment, $[H(T) - H_0]$; entropy, $S(T)$; and heat capacity at constant pressure, $C_p(T)$:

$$\frac{[H - H_0]}{RT} = T \frac{\partial}{\partial T} \ln(q), \tag{8}$$

$$\frac{S}{R} = T \frac{\partial}{\partial T} \ln(q) + \ln(q), \tag{9}$$

$$\frac{C_p}{R} = T^2 \frac{\partial^2}{\partial T^2} \ln(q) + 2T \frac{\partial}{\partial T} \ln(q), \tag{10}$$

where the explicit temperature dependence is omitted for brevity. If, rather than using an analytical expression for $q(T)$, the partition function were to be obtained by classical integration (or, for that matter, by quantum summation of rovibronic levels), the most frequent approach to computing the thermophysical properties would be to find the first and second derivatives of $q(T)$ via numeric finite difference approximations, using suitable finite increments δT, and computing the partition function at temperature T, as well as at least four more points, $T - 2\delta T$, $T - \delta T$, $T + \delta T$, and $T + 2\delta T$. A significant complication in

this finite-difference approach is determination of δT; this process typically requires trial and error, since the increment has to be sufficiently small as to produce a fully converged result for the derivatives, but not so small as to create numerical problems (for example, resulting in $q(T) - q(T - \delta T)$ differences that are of the same order of magnitude or even smaller than the convergence error in $q(T)$ itself).

An alternative (and more elegant) approach is to develop analytical expressions for the derivatives. In the case of quantum internal partition functions $q^{\text{quantum}}(T)$, which are obtained by summation of discrete rovibronic levels, m of which are bound, it can be shown that the pertinent expressions are:[40]

$$\Sigma_0 = \sum_{i=0}^{m} g_i \, e^{-\beta(E_i - E_0)} \tag{11}$$

$$\Sigma_1 = \sum_{i=0}^{m} \beta(E_i - E_0) \, g_i \, e^{-\beta(E_i - E_0)} \tag{12}$$

$$\Sigma_2 = \sum_{i=0}^{m} \left[\beta(E_i - E_0)\right]^2 \, g_i \, e^{-\beta(E_i - E_0)} \tag{13}$$

$$q^{\text{quantum}} = \Sigma_0 \tag{14}$$

$$\frac{[H - H_0]^{\text{quantum}}}{RT} = \frac{\langle E \rangle}{RT} = \frac{\Sigma_1}{\Sigma_0}, \tag{15}$$

$$\frac{S^{\text{quantum}}}{R} = \frac{\langle E \rangle}{RT} + \ln(q^{\text{quantum}}) = \frac{\Sigma_1}{\Sigma_0} + \ln(\Sigma_0) \tag{16}$$

$$\frac{C_p^{\text{quantum}}}{R} = \frac{\langle E^2 \rangle}{(RT)^2} - \left(\frac{\langle E \rangle}{RT}\right)^2 = \frac{\Sigma_2}{\Sigma_0} - \left(\frac{\Sigma_1}{\Sigma_0}\right)^2, \tag{17}$$

where E_0 is the zero point energy, E_i and g_i are the energy and degeneracy of the i^{th} rovibronic level, Σ_0 is the sum of Boltzmann factors over all $m + 1$ bound rovibronic levels, and Σ_1 and Σ_2 are sums similar to Σ_0, but each term in the sum is additionally weighted by the energy and by the square of the energy (in units of $k_B T$) of each level, respectively. Note also that the ratio Σ_1/Σ_0 corresponds to the average internal energy $\langle E \rangle/(RT)$, while Σ_2/Σ_0 corresponds to the average of the square of the internal energy $\langle E^2 \rangle/(RT)^2$. The SI includes the special case of representing n degrees of internal motion by n uncoupled

quantum harmonic oscillators (HO).

In contrast, if the partition function is obtained not by summation of quantum levels but rather via classical integration of the phase space over n degrees of freedom, as in Equation (6), then the resulting analytical expressions for the contributions corresponding to those degrees of internal motion are:

$$I_0 = \int_{x_1} \int_{x_2} ... \int_{x_n} e^{-\beta\, V(x_1,x_2,...,x_n)}\, dx_1 dx_2 ... dx_n \tag{18}$$

$$I_1 = \int_{x_1} \int_{x_2} ... \int_{x_n} \beta\, V(x_1,x_2,...,x_n)\, e^{-\beta\, V(x_1,x_2,...,x_n)}\, dx_1 dx_2 ... dx_n \tag{19}$$

$$I_2 = \int_{x_1} \int_{x_2} ... \int_{x_n} [\beta\, V(x_1,x_2,...,x_n)]^2\ e^{-\beta\, V(x_1,x_2,...,x_n)}\, dx_1 dx_2 ... dx_n \tag{20}$$

$$q^{\text{classical}} = \frac{(2\pi\, m\, k_B T)^{n/2}}{h^n}\, I_0 \tag{21}$$

$$\frac{[H - H_0]^{\text{classical}}}{RT} = \frac{n}{2} + \frac{I_1}{I_0} \tag{22}$$

$$\frac{S^{\text{classical}}}{R} = \frac{n}{2} + \frac{I_1}{I_0} + \ln\left[\frac{(2\pi\, m\, k_B T)^{n/2}}{h^n}\, I_0\right] \tag{23}$$

$$\frac{C_p^{\text{classical}}}{R} = \frac{n}{2} + \frac{I_2}{I_0} - \left(\frac{I_1}{I_0}\right)^2 \tag{24}$$

where the x_i are the coordinates in the general case. The SI includes the special case of representing n degrees of internal motion by n uncoupled classical harmonic oscillators (HO).

Finally, the classical partition function $q^{\text{classical}}(T)$ obtained via Equation (6) (or equivalently Equation (21)) is multiplied by the ratio of the quantum partition function to the classical partition function for the harmonic oscillator model, $q_{\text{HO}}^{\text{quantum}}/q_{\text{HO}}^{\text{classical}}$. This "semi-classical" correction, frequently referred to as the Pitzer-Gwinn appoximation,[41,42] is described in the SI, which also includes equations for the quantum corrected thermophysical properties.

As stated in the Introduction, the ultimate goal in determining the partition function

is to convert the derived thermophysical properties into a Gibbs free energy as a function of temperature: $G^\circ(T) = H^\circ(T) - TS^\circ(T)$ (the superscript indicates that the properties are at standard-state, as is required to calculate a reaction equilibrium constant, K_{eq}). In order to compute the free energy, we also need a reference heat of formation, $\Delta_f H^\circ$(298K). The approach to determining $\Delta_f H^\circ$(298K) for H on Cu(111) is based upon the methodology outlined in Ref. 43 and 44. A reference binding energy of -2.43 eV was taken from literature.[45–47] See SI for further details.

Results and Discussion

Surrogate Potential Energy Surface Construction

The surrogate PES construction for the PBE-D3(ABC) data includes 14,394 points as a combination of multivariate normal samples centered at the fcc site, Sobol sequence, as well as samples at the hcp, bridge and atop sites, as described in the Methods section. The one sample at the global minimum (fcc) was omitted from training in order for the transformation in Eq. (2) to be well-defined. Figure 3 demonstrates the surrogate energy evaluated at z-value corresponding to the fcc site, as a function of x and y coordinates. This 2D slice in potential energy space nicely illustrates the limitations of the two common analytical approaches. The local PES around the fcc is clearly neither harmonic nor flat, thereby confirming that neither the HO nor the FT model is appropriate, respectively. Instead, there is a broad range of energetically accessible area, confined within the repulsive walls of the atop sites.

In principle, for computing thermodynamic properties in the canonical ensemble, the global accuracy of the MP-NN surrogate PES is not as important as the accuracy of the approximation integrand of the Boltzmann factor, $e^{-E/kT}$, since the main target of interest is the partition function evaluation, *i.e.* the integration in Eq. (6). Figure 4 demonstrates the surrogate PES accuracy for various values of T, showing an overall relative root-mean-square error (RMSE) under 8% across a range of T. Note the overall accuracy improvement

with increasing T: this is due to the fact that the integrand in Eq. (6) becomes less sharp near its maxima (*i.e.*, near energy minima), and small errors in the surrogate PES are less significant.

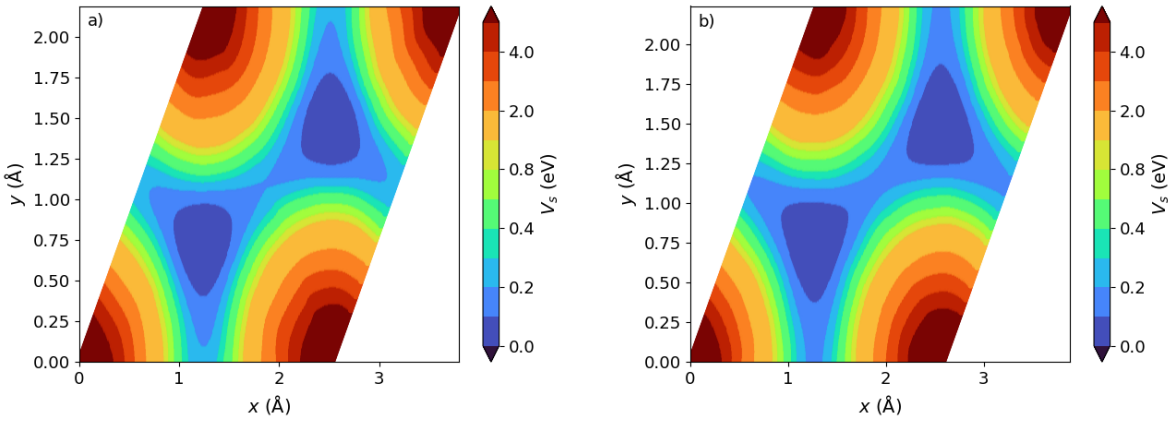

Figure 3: The MP-NN surrogate PES evaluated at the z-slice corresponding to the fcc minimum of a) surrogate 1, which was obtained using a single MVN set with the PBE-D3(ABC) functional, and b) surrogate 2, which was obtained using two MVN sets with the BEEF-vdW functional.

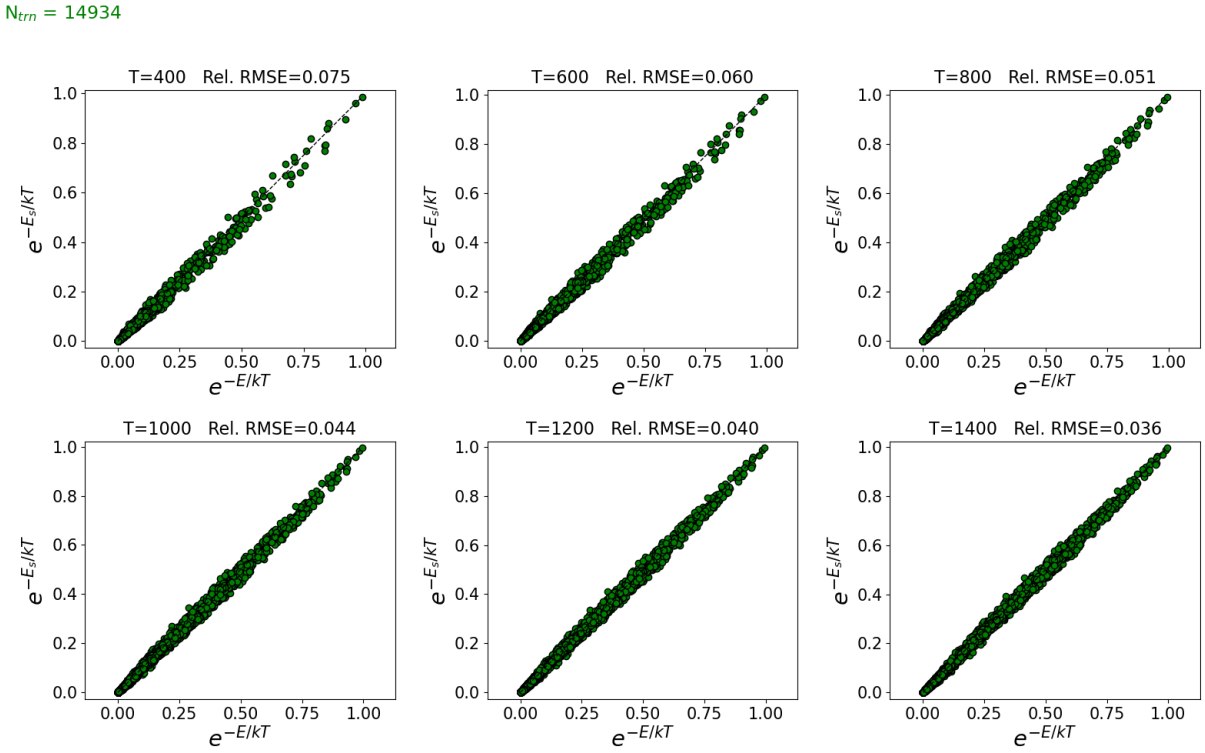

Figure 4: Partition function integrand approximation for surrogate 1, which was obtained using a single MVN sampling set (with the PBE-D3(ABC) functional), for a range of T values.

For the BEEF-vdW functional, where two sets of MVN samples (centered at both minima) were included, in addition to a Sobol sequence and samples at the bridge and atop sites. Here we removed two samples, corresponding to the exact fcc and hcp sites, in order for the transformation in Eq. (2) to be well-defined and not harm the training accuracy. Figure 5 shows that similar to the PBE-D3(ABC) case, the BEEF-vdW-based surrogate PES tends to be more accurate with larger values of T. The overall performance of the BEEF-vdW surrogate is better, even though for the BEEF-vdW functional we used fewer DFT evaluations (N=9,380) compared to the PBE-D3(ABC) case (N=14,934). We attribute this superiority to the improved sampling scheme: the BEEF-vdW functional was sampled according to an MVN near both local minima, fcc and hcp sites, whereas the PBE-D3(ABC) functional only accounted for the fcc site.

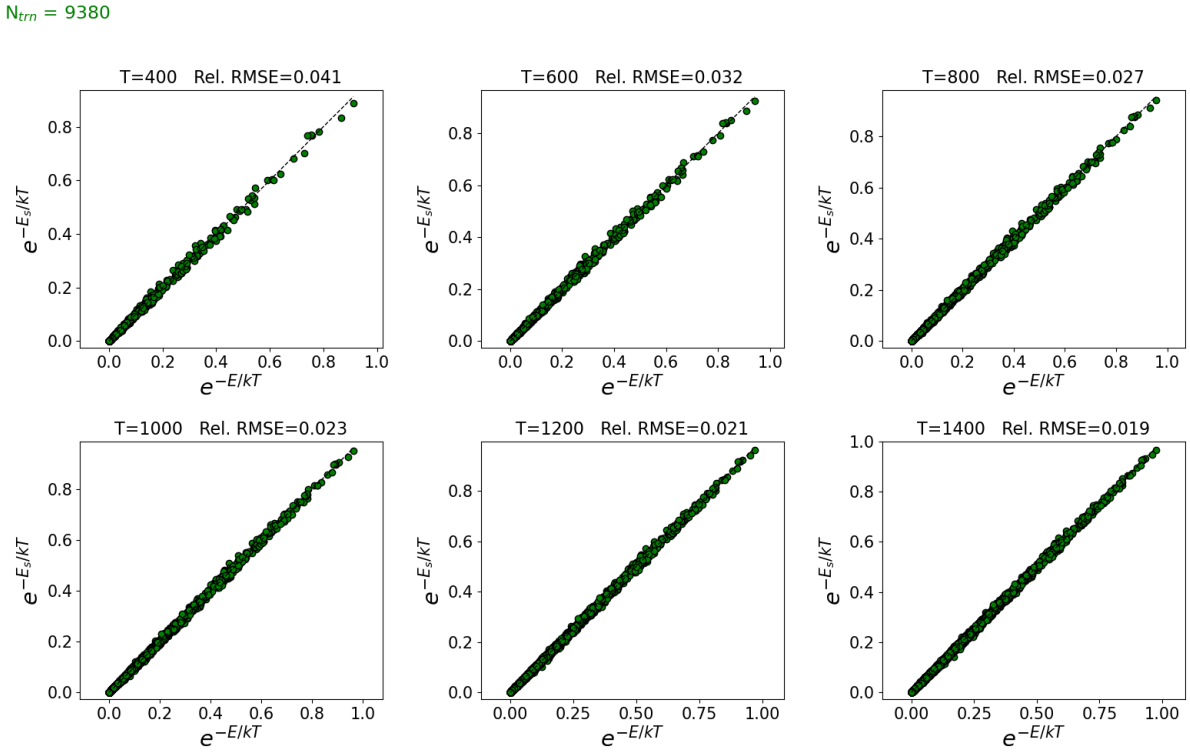

Figure 5: Partition function integrand approximation for surrogate 2, which was obtained using two MVN sampling sets (with the BEEF-vdW functional), for a range of T values.

Thermophysical Properties

The NITROGEN[39] program interfaced with the PBE-D3(ABC) based MP-NN surrogate PES yielded the wavenumbers 880.9 cm^{-1}, 881.1 cm^{-1} and 1084.8 cm^{-1} for the H atom in the fcc position. Interfaced with the BEEF-vdW based surrogate PES the program yielded the wavenumbers 793.674 cm^{-1}, 795.640 cm^{-1} and 1076.139 cm^{-1}. These wavenumbers were used in the HO model as well as for the z degree of freedom in the HT and FT partition functions. The resulting DVR partition function is plotted in Figure 6 (black stars).

These DVR results represent benchmark values for the partition function and thence thermophysical properties of H on Cu(111). Note: in referring to these results as "benchmark" for thermophysical properties, we are not asserting that the DFT functionals are themselves of benchmark accuracy in terms of the binding energy at 0K; rather, we mean that, given an underlying potential energy surface, the DVR results are an accurate benchmark for

comparison with methods of obtaining partition functions (*e.g.* hindered translator, phase space integration, and other methods).

The phase space integration over each MP-NN surrogate PES was performed at 31 evenly spaced temperature values between 290–1400 K. To satisfy the convergence criterion $c < 1.0\times10^{-6}$ (Equation 7) for all T, surrogate 1 required 4,848,591 geometries in the sampling space with the primitive unit cell area (1/9th of a 3×3 unit cell area), which required a run time of 2h 1m 46s on 1 CPU core. For surrogate 2, 4,843,036 geometries were required, and the integration run time was 3h 10m 34s on 1 CPU core. An integration convergence study is included in the Supporting Information. Note that the PSI calculation using surrogate 2 required a slightly longer run time. The longer run time is expected, since surrogate 2 used 10x more units per hidden layer, even though it used fewer training samples. The training samples, *i.e.* the actual DFT evaluations are orders of magnitude more expensive compared to the time required for the NN training and its evaluation to achieve integration convergence. Therefore, the computational efficiency is almost exclusively determined by the amount of DFT single-point calculations required to generate the training data. As such, the surrogate 2 approach of sampling MVN near *both* minima is computationally more efficient, as it leads to a more accurate surrogate PES with fewer samples. Full convergence study of the surrogate as the number of samples grows is provided in the Supporting Information.

The results of the PSI method are shown in Figure 6 (solid black line). Also included in Figure 6 are the two common analytical models for the limiting cases, harmonic oscillator (red dash-dot line) and free translator (blue dashed line), as well as the hindered translator model of Sprowl et al. (yellowed dotted line). It should be noted that when calculating the FT partition function (which is proportional to the area), the partition function would not actually be confined to a single unit cell, but rather it would be calculated assuming a much larger area (for example, using $V^{2/3}$, analogous to 3D gas volume at 1 bar), as it only holds when the translational energy levels are separated by energies much smaller than the available thermal energy (kT), as discussed in Ref 8. This larger area would lead to a

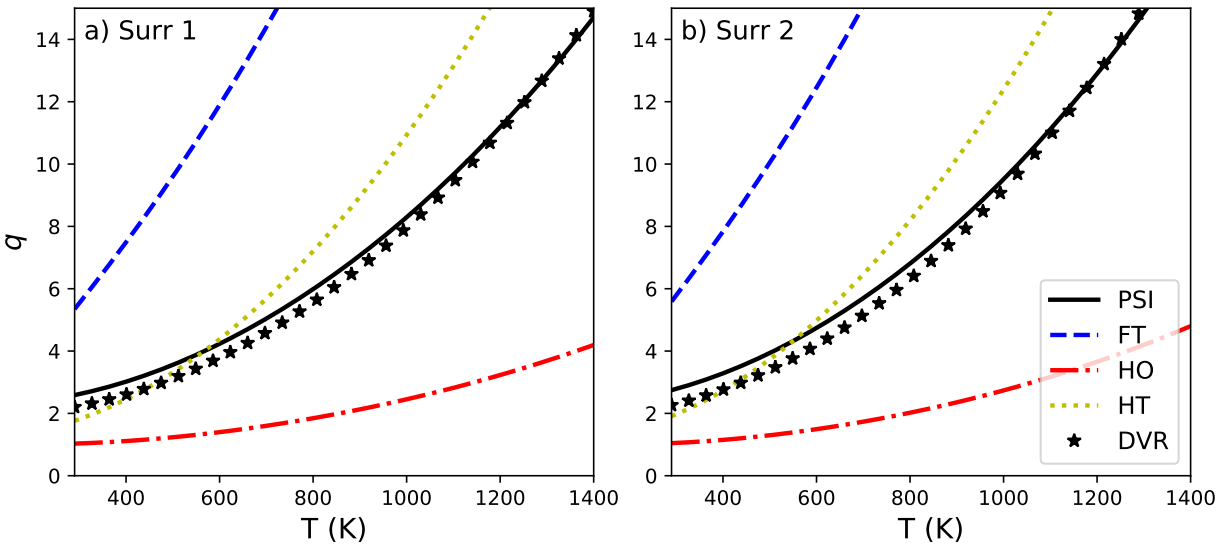

Figure 6: The partition function per cell for a hydrogen atom on Cu(111) at 290-1400 K, evaluated using the phase space integration (PSI) of the MP-NN surrogate PES, the free translator (FT), the 3D harmonic oscillator (HO), the hindered translator (HT), and discrete variable representation (DVR). a) shows the results for surrogate 1, which was obtained with a single MVN set, and b) shows the results for surrogate 2, which was obtained with two MVN sets.

substantially larger partition function. To facilitate a more direct comparison, however, we chose to use the same unit cell area for the FT, HT, PSI and DVR models (the HO model is independent of area).

The hindered translator model is qualitatively correct over the temperature range of interest. It is worth noting that the hindered translator model includes a correction for the zero-point energy contribution due to overestimation at lower temperatures as discussed in Ref. 8, but it is not important at the temperature range applied here (>300 K). Indeed, at the lower temperatures, the simpler model is in quantitative agreement with the benchmark DVR calculations. At higher temperatures, however, it significantly overpredicts the partition function. The PSI method, in contrast, is in excellent agreement with the DVR results over the entire temperature range. In all cases, we observe that the PSI method is between the tight-binding limit (HO) and weak binding limit (FT), and that the adsorbate becomes increasingly anharmonic as the temperature increases.

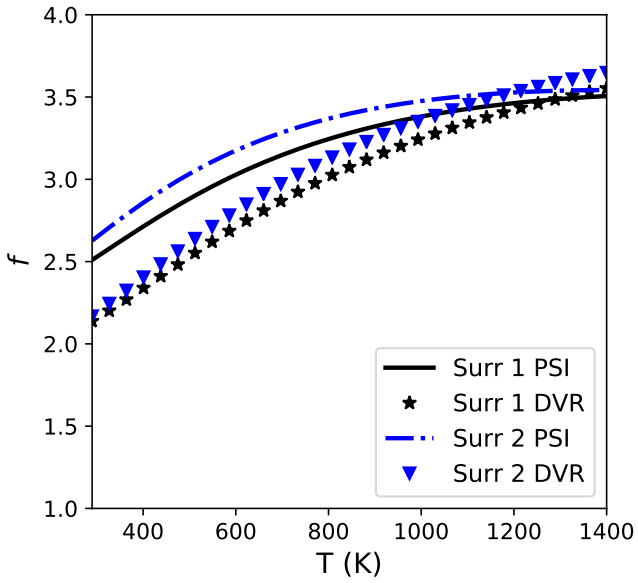

Figure 7: The anharmonic correction factor evaluated with PSI and direct state counting for both surrogate PES. A value of f=1 would be perfectly harmonic.

Since the standard approach to computing the partition function of H on Cu(111) typically would be to use the harmonic oscillator model, it is useful to consider how the DVR and PSI methods compare to the pure HO result. The anharmonic correction factor, $f \equiv q/q_{\mathrm{HO}}$, is plotted in Figure 7. Qualitatively, we observe that the PSI slightly overpredicts the anharmonicity at lower temperatures, and it slightly underpredicts the anharmonicity at higher temperatures. Overall, however, the discrepancy between PSI and DVR is small, and the PSI approach is in excellent agreement with the benchmark data.

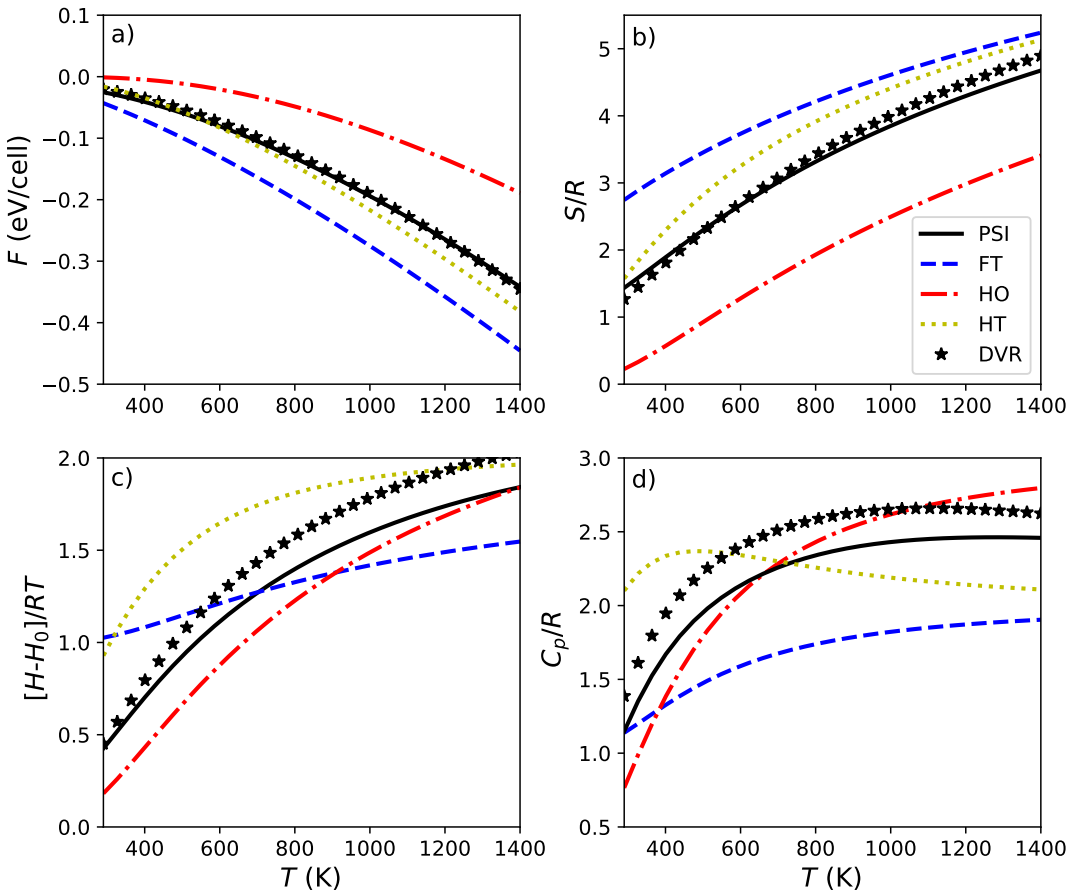

Figure 8: a) Helmholtz free energy F, b) entropy S/R, c) enthalpy increment $(H(T) - H(0))/RT$, and d) heat capacity C_p/R, versus temperature for surrogate 2.

Figure 8 presents the Helmholtz free energy, $F = -k_B T \ln(q)$, as well as the three thermophysical properties, derived from the partition function evaluated at 1/9th monolayer coverage: 3D translational entropy (no concentration-dependent terms included), S/R; enthalpy increment, $[H(T) - H(0)]/RT$; and heat capacity at constant pressure, C_p/R, for surrogate 2. Since both surrogates give qualitatively the same model profiles, as demonstrated in Figure 6, we only present the thermophysical properties for the surrogate constructed from the more feasible sampling method here. An analogous figure for surrogate 1 can be found in the Supporting Information. Additionally, the Supporting Information contains a comparison to finite difference results at various finite increments δT. Of particular interest is the Figure 8b, the entropy S/R. Here we observe that the anharmonic PSI (along with

the benchmark DVR) results are effectively halfway between the lower (HO) and upper (FT) limits, and slightly lower than predicted by the HT model. For the enthalpy increment, Figure 8c, the HO model is too low but otherwise predicts the correct temperature dependence, whereas the FT model significantly underestimates the temperature dependence. The PSI approach slightly underpredicts the enthalpy increment and heat capacity, but is overall in better agreement with the DVR results than the HO and the FT models.

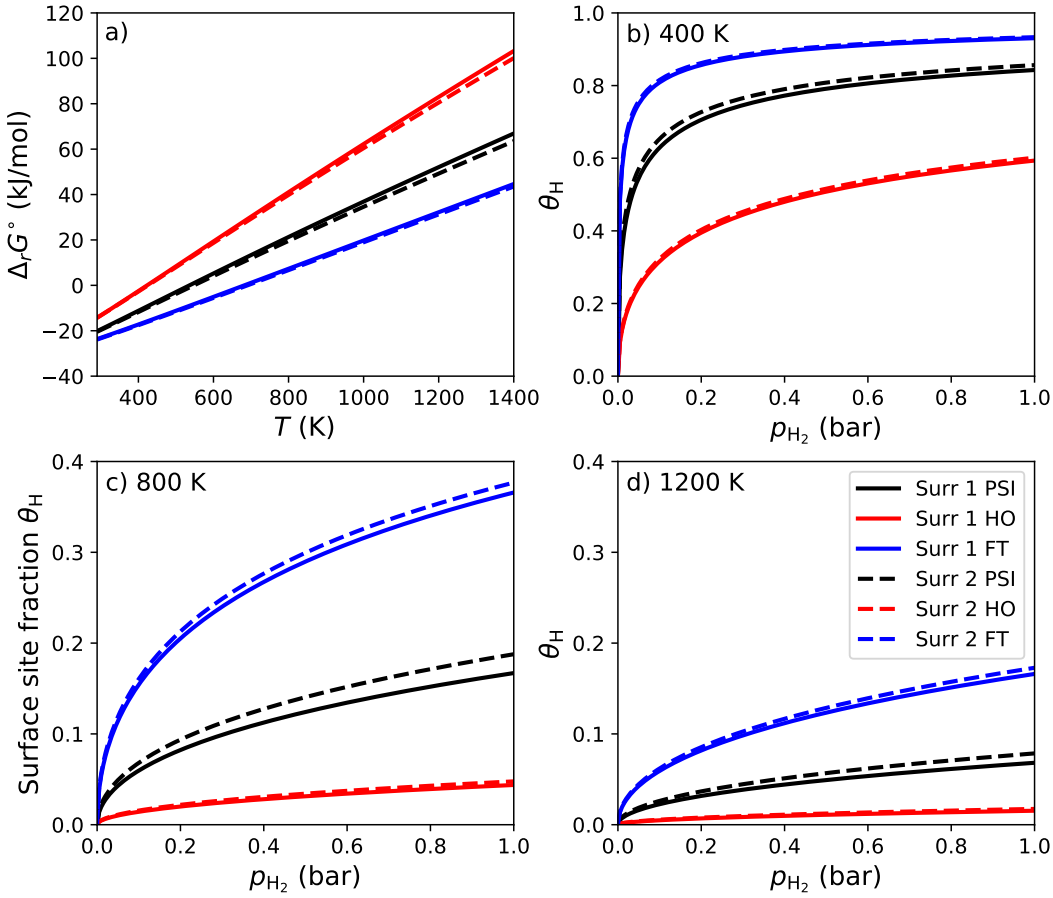

Figure 9: a) The standard-state Gibbs free energy of reaction $\Delta_r G^\circ$ of dissociative adsorption of H_2, obtained with PSI of each surrogate PES, compared with the corresponding HO and FT models. The remaining three subplots show the Langmuir adsorption isotherms obtained using PSI, the HO and the FT approximations on both surrogates at b) 400 K, c) 800 K, and d) 1200 K.

Of perhaps greater interest than the derived thermophysical properties is the Gibbs free energy. We look at the Gibbs free energy of reaction of the dissociative adsorption of H_2,

$H_2 + 2* \rightarrow 2H*$. In order to calculate the standard-state Gibbs free energy of reaction (the standard state is necessary in relation to the reaction equilibrium constant), we need to define a standard-state surface concentration. There is currently no universally accepted standard state for adsorbates, but the surface site fraction θ°=0.5 is commonly used and proposed as a standard.[48] We choose follow that convention for the standard state here, since it is recommended for the Langmuir isotherm for a homolytic dissociative adsorption[48] which we apply and plot in Figure 9. With a standard state of θ°=0.5 defined, we can calculate the standard-state entropy, S°, the concentration dependence of which is slightly different depending on the model in question.[49] This effect is discussed further in the SI.

The standard-state Gibbs free energy of reaction for the dissociative adsorption of H_2 is plotted in Figure 9. Because the PSI approach correctly predicts a greater value of entropy for the adsorbed hydrogen atom, the corresponding free energy is consistently lower than predicted by the HO approximation at all temperatures. Additionally, we observe that the two functionals (and thus also the two MP-NN models) are in excellent agreement. The fact that the two functionals are effectively indistinguishable suggests that they predict similar curvature of the 3D potential energy surface. However, it is worth emphasizing that both functionals used the same literature binding energy of -2.43 eV. If this binding energy were taken from the corresponding DFT calculation instead, then the offset between the PBE-D3(ABC) (surrogate 1) and the BEEF-vdW (surrogate 2) results would increase, reflecting the additional difference in the binding energies. Collectively, the plots in Figures 8 and the Gibbs free energy of reaction in Figure 9 show that the PSI approach gives thermophysical properties and free energies that are substantially different than both commonly used limiting cases. These results indicate significant effects on theoretically estimated thermodynamics and rates of reactions involving H on Cu(111), especially reactions that involve both gas-phase and surface species, where no cancellation of anharmonicity will occur.

Finally, with the free energy for dissociative adsorption of H_2, we can quantify how the choice of the partition function for H on Cu(111) influences the resulting coverage in a mi-

crokinetic model. Figure 9 illustrates the surface site fraction of H* using a simple Langmuir isotherm: $\theta_{\mathrm{H}} = \sqrt{K_{eq} p_{\mathrm{H}_2}} / \left(1 + \sqrt{K_{eq} p_{\mathrm{H}_2}}\right)$, where K_{eq} is the equilibrium constant for dissociative adsorption, and p_{H_2} is the partial pressure of hydrogen gas. While the Langmuir isotherm we use here describes adsorption adequately for many cases,[48] more accurate, model specific coverage equations could be obtained from equating chemical potentials as described in Ref. 49. It is evident that the increased entropy for H* (relative to the HO model) directly corresponds to higher coverage for a given temperature and pressure. In fact, for the ranges of temperature and pressures considered, the PSI predicts on average a surface site fraction that is 1.6–4.6 times as large as the HO model.

Conclusions

A new method is presented for creating surrogate potential energy surfaces for the translation of an adsorbate on a catalytic surface. This method is designed to preserve accuracy nearest to the global and local minima. This surrogate potential energy surface is used within discrete variable representation calculations, which serve as benchmark values, as well as classical phase space integration methods. The DVR and PSI methods are in excellent quantitative agreement over the entire temperature range of interest. Together, the two models confirm that the motion of H on Cu(111) is highly anharmonic. Conventional models, such as the harmonic oscillator, will severely underpredict the entropy of adsorbates on catalytic surfaces. This underprediction of adsorbate entropy will adversely impact microkinetic mechanisms, leading to potentially lower surface coverage.

The present calculations all were performed in the low-coverage limit. Lateral interactions were not considered. However, the approach presented in this work in principle could be extended to include coverage dependence. Anharmonic, coverage-dependent partition functions, building on this approach, are underway. Additionally, the current approach is being extended to include polyatomic adsorbates. In addition to the three degrees of freedom

for frustrated translation along the surface, two and three dimensional frustrated rotational motion will be considered for linear and nonlinear adsorbates, respectively.

This work is part of a broader effort to provide accurate microkinetic mechanisms within an automated computational chemistry framework. The minima-preserving neural network and the phase space integration routines are provided as part of the AdTherm package.[50] The method is to be a part of the Sandia National Laboratories' Exascale Catalytic Chemistry (ECC) computational framework for automated chemistry.[51] The results of these calculations will be incorporated into the Active Thermochemical Tables (ATcT) of Argonne National Laboratory[52] as it expands to include thermophysical properties for adsorbates. Finally, the improved thermodynamic properties will be included both in the database and in the group rules used by the automatic microkinetics software for heterogeneous catalysis, RMG.[43,53–56]

Acknowledgement

The authors wish to acknowledge many helpful discussions with Eric Hermes, Habib Najm, and Judit Zádor at Sandia National Laboratories, Bryan Changala at Harvard-Smithsonian Center for Astrophysics, and Charles Campbell at the University of Washington. This work was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences, under Contract No. DE-AC02-06CH11357 (ANL) through the Computational Chemical Sciences Program (D.H.B.) and the Gas-Phase Chemical Physics Program (B.R.). KB and CFG are supported by Grant No. 0000232253. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Supporting Information Available

The following materials are available free of charge in the Supporting Information:

- Thermophysical property equations for the quantum and classical harmonic oscillators
- Pitzer-Gwinn correction details
- Gibbs free energy and Langmuir adsorption isotherm calculation details
- wavenumber table for H in the Cu(111) fcc site
- surrogate convergence study
- integration convergence study
- hindered translator equation
- thermodynamic properties for surrogate 1
- tabulated thermodynamic properties
- comparison of thermodynamic property derivation methods.

References

(1) Motagamwala, A. H.; Dumesic, J. A. Microkinetic Modeling: A Tool for Rational Catalyst Design. *Chemical Reviews* **2021**, *121*, 1049–1076.

(2) Mhadeshwar, A. B.; Wang, H.; Vlachos, D. G. Thermodynamic Consistency in Microkinetic Development of Surface Reaction Mechanisms. *The Journal of Physical Chemistry B* **2003**, *107*, 12721–12733.

(3) Chorkendorff, I.; Niemantsverdriet, J. W. *Concepts in Modern Catalysis and Kinetics*; Wiley: Weinheim, 2007.

(4) Campbell, C. T.; Sellers, J. R. V. The Entropies of Adsorbed Molecules. *Journal of the American Chemical Society* **2012**, *134*, 18109–18115.

(5) Savara, A.; Schmidt, C. M.; Geiger, F. M.; Weitz, E. Adsorption Entropies and Enthalpies and Their Implications for Adsorbate Dynamics. *The Journal of Physical Chemistry C* **2009**, *113*, 2806–2815.

(6) Collinge, G.; Yuk, S. F.; Nguyen, M.-T.; Lee, M.-S.; Glezakou, V.-A.; Rousseau, R. Effect of Collective Dynamics and Anharmonicity on Entropy in Heterogenous Catalysis: Building the Case for Advanced Molecular Simulations. *ACS Catalysis* **2020**, *10*, 9236–9260.

(7) Amsler, J.; Plessow, P. N.; Studt, F.; Bučko, T. Anharmonic Correction to Adsorption Free Energy from DFT-Based MD Using Thermodynamic Integration. *Journal of Chemical Theory and Computation* **2021**, *17*, 1155–1169.

(8) Sprowl, L. H.; Campbell, C. T.; Árnadottir, L. Hindered Translator and Hindered Rotor Models for Adsorbates: Partition Functions and Entropies. *The Journal of Physical Chemistry C* **2016**, *120*, 9719–9731.

(9) Jørgensen, M.; Grönbeck, H. Adsorbate Entropies with Complete Potential Energy Sampling in Microkinetic Modeling. *The Journal of Physical Chemistry C* **2017**, *121*, 7199–7207.

(10) Jørgensen, M.; Chen, L.; Grönbeck, H. Monte Carlo Potential Energy Sampling for Molecular Entropy in Zeolites. *The Journal of Physical Chemistry C* **2018**, *122*, 20351–20357.

(11) Bajpai, A. H.; Mehta, P.; Frey, K.; Lehmer, A. M.; Schneider, W. F. Benchmark First-Principles Calculations of Adsorbate Free Energies. *ACS Catalysis* **2018**, *8*, 1945–1954.

(12) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Physical Review Letters* **2007**, *98*, 146401.

(13) Lym, J.; Gu, G. H.; Jung, Y.; Vlachos, D. G. Lattice Convolutional Neural Network Modeling of Adsorbate Coverage Effects. *The Journal of Physical Chemistry C* **2019**, *123*, 18951– 18959.

(14) Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C. Machine Learning for Molecular Simulation. *Annual Review of Physical Chemistry* **2020**, *71*, 361–390.

(15) Miksch, A. M.; Morawietz, T.; Kästner, J.; Urban, A.; Artrith, N. Strategies for the Construction of Machine-Learning Potentials for Accurate and Efficient Atomic-Scale Simulations. *Machine Learning: Science and Technology* **2021**, *2*, 031001.

(16) Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chemical Reviews (in press)* **2021**, doi:10.1021/acs.chemrev.0c00868.

(17) Bross, D. H.; Jasper, A. W.; Ruscic, B.; Wagner, A. F. Toward accurate high temperature anharmonic partition functions. *Proceedings of the Combustion Institute* **2019**, *37*, 315–322.

(18) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I. et al. Quantum ESPRESSO: a modular and open-source software project for quantum simulations of materials. *Journal of Physics: Condensed Matter* **2009**, *21*, 395502.

(19) Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Buongiorno Nardelli, M.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M. et al. Advanced capabili-

ties for materials modelling with Quantum ESPRESSO. *Journal of Physics: Condensed Matter* **2017**, *29*, 465901.

(20) Quantum ESPRESSO. `http://www.quantum-espresso.org/`, v. 6.4 (for surr 1 with PBE-D3(ABC)) and v. 6.6 (for surr 2 with BEEF-vdW).

(21) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77*, 3865–3868.

(22) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements of H-Pu. *The Journal of Chemical Physics* **2010**, *132*, 154104.

(23) Axilrod, B. M.; Teller, E. Interaction of the van der Waals Type Between Three Atoms. *The Journal of Chemical Physics* **1943**, *11*, 299.

(24) Muto, Y. On the forces acting between non-polar molecules. *Proceedings of the Physico-Mathematical Society of Japan* **1943**, *17*, 629.

(25) Dal Corso, A. Pseudopotentials periodic table: From H to Pu. *Computational Materials Science* **2014**, *95*, 337–350.

(26) pslibrary by A. Dal Corso. 2013; We used the pseudopotentials H.pbe-kjpaw_psl.1.0.0.UPF and Cu.pbe-dn-kjpaw_psl.1.0.0.UPF from https://dalcorso.github.io/pslibrary/, accessed: 2019-08-01.

(27) Bahn, S. R.; Jacobsen, W., K An object-oriented scripting interface to a legacy electronic structure code. *Computing in Science & Engineering* **2002**, *4*, 56–66.

(28) Marzari, N.; D., V.; De Vita, A.; Payne, M. C. Thermal Contraction and Disordering of the Al(110) Surface. *Physical Review Letters* **1999**, *82*, 3296.

(29) Sobol, I. M. Distribution of Points in a Cube and Approximate Evaluation of Integrals. *USSR Computational Mathematics and Mathematical Physics* **1967**, *7*, 86–112.

(30) Wellendorff, J.; Lundgaard, K. T.; Mogelhoj, 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. *Physical Review B* **2012**, *85*.

(31) Thonhauser, T.; Zuluaga, S.; Arter, C. A.; Berland, K.; Schröder, E.; Hyldgaard, P. Spin Signature of Nonlocal Correlation Binding in Metal-Organic Frameworks. *Physical Review Letters* **2015**, *115*, 136402.

(32) Thonhauser, T.; Cooper, V. R.; Li, C.; Puzder, A.; ; Hyldgaard, P.; Langreth, D. C. Van der Waals density functional: Self-consistent potential and the nature of the van der Waals bond. *Physical Review B* **2007**, *76*, 125112.

(33) libbeef by J. Voss. 2021; `https://github.com/vossjo/libbeef`, Accessed: 2020-11-09.

(34) Hamann, D. R. Optimized norm-conserving Vanderbilt pseudopotentials. *Physical Review B* **2013**, *88*, 085117.

(35) Schlipf, M.; Gygi, F. Optimization algorithm for the generation of ONCV pseudopotentials. *Computer Physics Communications* **2015**, *196*, 36–44.

(36) SG15 ONCV Potentials by D. R. Hamann. 2013; We used the pseudopotentials H_ONCV_PBE-1.2.upf and Cu_ONCV_PBE-1.2.upf from http://www.quantum-simulation.org/potentials/sg15_oncv/, accessed: 2020-12-16.

(37) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. ArXiv (cs.LG), 30 Jan 2017, 1412.6980v9 (accessed 2021-08-20).

(38) Colbert, D. T.; Miller, W. H. A Novel Discrete Variable Representation for Quantum Mechanical Reactive Scattering via the S-Matrix Kohn Matrix. *J. Chem. Phys.* **1992**, *96*, 1982–1991.

(39) NITROGEN, Numerical and Iterative Techniques for Rovibronic Energies with General Internal Coordinates, a program by P. B. Changala, http://www.colorado.edu/nitrogen. {https://sourceforge.net/projects/nitrogen-downloads/files/}, Accessed: 2020-11-06.

(40) Ruscic, B.; Bross, D. H. Thermochemistry. *Computer-Aided Chemical Engineering* **2019**, *45*, 3–114.

(41) Pitzer, K. S.; Gwinn, W. D. Energy Levels and Thermodynamic Functions for Molecules with Internal Rotation I. Rigid Frame with Attached Tops. *The Journal of Chemical Physics* **1942**, *10*, 428.

(42) Isaacson, A. D.; Truhlar, D. G. The Accuracy of the Pitzer-Gwinn Method for Partition Functions of Anharmonic Vibrational Modes. *The Journal of Chemical Physics* **1942**, *75*, 4090.

(43) Blondal, K.; Jelic, J.; Mazeau, E.; Studt, F.; West, R. H.; Goldsmith, C. F. Computer-Generated Kinetics for Coupled Heterogeneous/Homogeneous Systems: A Case Study in Catalytic Combustion of Methane on Platinum. *Industrial & Engineering Chemistry Research* **2019**, *58*, 17682–17691.

(44) Klippenstein, S. J.; Harding, L. B.; Ruscic, B. Ab initio Computations and Active Thermochemical Tables Hand in Hand: Heats of Formation of Core Combustion Species. *The Journal of Physical Chemistry A* **2017**, *121*, 6580–6602.

(45) Xu, L.; Lin, J.; Bai, Y.; Mavrikakis, M. Atomic and Molecular Adsorption on Cu(111). *Topics in Catalysis* **2018**, *61*, 736–750.

(46) Ertl, G. *The Nature of the surface chemical bond*; North-Holland Publishing Company, 1979.

(47) Gaydon, A. G. *Dissociation energies and spectra of diatomic molecules*; Chapman and Hall, 1968.

(48) Savara, A. Standard States for Adsorption on Solid Surfaces: 2D Gases, Surface Liquids, and Langmuir Adsorbates. *The Journal of Physical Chemistry C* **2013**, *117*, 15710–15715.

(49) Campbell, C. T.; Sprowl, L. H.; Árnadóttir, L. Equilibrium Constants and Rate Constants for Adsorbates: Two-Dimensional (2D) Ideal Gas, 2D Ideal Lattice Gas, and Ideal Hindered Translator Models. *The Journal of Physical Chemistry C* **2016**, *120*, 10283–10297.

(50) AdTherm. `https://github.com/franklingoldsmith/adtherm`, v. 0.1.

(51) Exascale Catalytic Chemistry (ECC) Project at Sandia National Laboratories. `https://ecc-project.sandia.gov/`, Accessed: 2021-08-23.

(52) Ruscic, B.; Pinzon, R. E.; Morton, M. L.; von Laszevski, G.; Bittner, S. J.; Nijsure, S. G.; Amin, K. A.; Minkoff, M.; Wagner, A. F. Introduction to Active Thermochemical Tables: Several "Key" Enthalpies of Formation Revisited. *The Journal of Physical Chemistry A* **2004**, *108*, 9979–9997.

(53) Goldsmith, C. F.; West, R. Automatic Generation of Microkinetic Mechanisms in Heterogeneous Catalysis. *The Journal of Physical Chemistry C* **2017**, *121*, 9970–9981.

(54) Liu, M.; Dana, A. G.; Johnson, M.; Goldman, M.; Jocher, A.; Payne, A. M.; Grambow, C.; Han, K.; Yee, N. W.-W.; Mazeau, E. et al. Reaction Mechanism Generator v3.0: Advances in Automatic Mechanism Generation. *Journal of Chemical Information and Modeling* **2021**, *61*, 2686–2696.

(55) Mazeau, E.; Satpute, P.; Blöndal, K.; Goldsmith, F.; West, R. Automated Mechanism

Generation Using Linear Scaling Relationships and Sensitivity Analyses Applied to Catalytic Partial Oxidation of Methane. *ACS Catalysis* **2021**, *11*, 7114–7125.

(56) Kreitz, B.; Sargsyan, K.; Mazeau, E. J.; Blöndal, K.; West, R. H.; Wehinger, G. D.; Turek, T.; Goldsmith, C. F. Quantifying the Impact of Parametric Uncertainty on Automatic Mechanism Generation for CO_2 Hydrogenation on Ni(111). *JACS Au (in press)* **2021**, doi:10.1021/jacsau.1c00276.

Graphical TOC Entry