THE JOURNAL OF PHYSICAL CHEMISTRY C

pubs.acs.org/JPCC Article

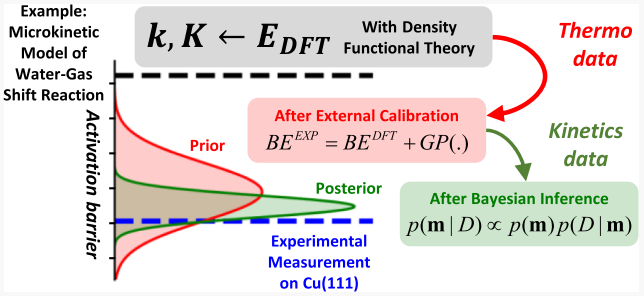
Leveraging Thermochemistry Data to Build Accurate Microkinetic Models

Huijie Tian and Srinivas Rangarajan*

Cite This: J. Phys. Chem. C 2020, 124, 5740−5748 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Ab initio microkinetic modeling, parameterized using density functional theory (DFT) energies, is a common tool to quantify reaction rates and analyze reaction mechanisms a priori in heterogeneous catalysis. Such models, however, often have large prediction errors even if they include plausible reaction steps and correctly model the active sites; this is partially due to the intrinsic inaccuracies of the chosen DFT functional. Borrowing concepts from Bayesian calibration theory, we show that transferable data-driven corrections to DFT energies in the form of Gaussian process models trained on single-crystal adsorption calorimetry data can improve the accuracy of microkinetic models substantially. Specifically, we demonstrate that such corrections improve the predictive accuracy of the microkinetic model of the water-gas shift reaction on single-crystal Cu(111) surface by 3 orders of magnitude. We finally show that Gaussian process corrections serve as informed priors in a Bayesian experimental design framework to learn an accurate a posteriori microkinetic model from few kinetic experiments. We posit that these results suggest that even infusing small, related, high-fidelity thermochemistry data, when available, can systematically and substantially improve the predictive accuracy of microkinetic models.



INTRODUCTION

Mean-field microkinetic modeling is a common tool to quantify the time evolution of surface intermediates and reaction fluxes in heterogeneous catalytic systems.[1] When parameterized by first-principles calculations, such as density functional theory (DFT), microkinetic modeling provides a predictive tool to link the atomistic information (energies, activation barriers) with microscopic/macroscopic observables (reaction rate, activity, and selectivity) of a catalytic reaction system.[2,3] However, the predictions of ab initio microkinetic models often deviate from experimental observations in terms of the turn over frequency (TOF); indeed, the error could be several orders of magnitude.[2,4,5] It is hard then to attribute the discrepancy between microkinetic models and experimental observations to an incorrectly assumed active site, or missing elementary reactions in the mechanism, because the inputs to the microkinetic model (especially DFT-derived kinetic and thermochemical parameters) embed, within themselves, certain level of error.[2] A recent benchmarking study compares the binding energy from experimental [single-crystal adsorption microcalorimetry and temperature-programmed desorption (TPD)] measurements to DFT predictions.[6] It reports that six common gradient generalized approximation (GGA) functionals have 20−40 kJ/mol error on the binding energy of small molecules or intermediates adsorbed on transition metals. Such errors result in miscalculating kinetics or thermodynamic parameters by 2−4 (two to four) orders of magnitude at typical reaction temperatures of 500 K, which in turn usually results in large predictive errors in a microkinetic model that otherwise has captured the right site and chemistry. Consequently, bridging the gap between the ab initio microkinetic models and real kinetic experiments is still an open question in computational catalysis.[5,7]

The quality of microkinetic models is largely improved in two ways. First, the accuracy of kinetic parameters is improved by employing a higher level of theory[8,9] albeit at a greater computational cost that can become prohibitive as the complexity of the reaction system increases.[10,11] Second, the kinetic and thermochemical parameters can be refined in a rigorous way to fit the microkinetic model to experimental kinetic data.[4,5,7,12] This is a postdictive correction, thereby necessitating a certain, albeit modest, amount of experiments; however, the resulting model is robust, often shows good extrapolative power to new conditions, and can be used infer the reaction mechanism and guide further experiments, kinetics, or otherwise.

A third approach to address the inaccuracies in microkinetic models is to rigorously evaluate and report the associated

Received: January 17, 2020
Revised: February 14, 2020
Published: February 17, 2020

https://dx.doi.org/10.1021/acs.jpcc.0c00491
J. Phys. Chem. C 2020, 124, 5740−5748

prediction uncertainties by ascertaining a statistical distribution of the uncertainties on the kinetic and thermodynamic parameters.[13−15] Several techniques have been adopted to this end. The first is the Bayesian error estimation functional developed by the Nørskov group which is trained on surface adsorption dataset using Bayesian inference.[16] The resulting estimates of energies are therefore fairly accurate, especially for metal surfaces, and the functional also provides uncertainty estimates that can then be propagated to get uncertainty bands for the microkinetic model predictions.[14] The second method, proposed by the Heyden group, uses energies estimated from a collection of theories/functionals to build a statistical distribution which can also be propagated further to get the variance of the microkinetic model predictions.[15] The third method is to use the uncertainties in the parameters of a data-driven model that is trained on DFT energies as adopted by the Vlachos group. Once these uncertainties are known, they can serve as priors to further refine the parameters through a Bayesian calibration process.[5,17]

In this work, we propose a data-driven method that enables learning from thermochemistry data initially, and kinetic experiments subsequently, to improve microkinetic modeling predictions. The main illustrative catalytic system we consider here is the water-gas shift (WGS) reaction ($CO(g) + H_2O(g) \rightarrow CO_2(g) + H_2(g)$) catalyzed by Cu(111), for which we show that mean-field microkinetic model parameterized with DFT energies substantially underpredicts single-crystal kinetic experiments; we argue that the origin of this inaccuracy largely ought to arise from errors in DFT. We first show that it is possible to substantially improve microkinetic model predictions by simply correcting the DFT energies with a Gaussian process (GP) model[18] trained on the mismatch between single-crystal adsorption calorimetry data and its analogous DFT-computed binding energies of small adsorbates. As GP is a Bayesian nonparametric data-driven model that also quantifies the uncertainties in the response values, we then show that this information can be used as the prior to systematically identify the necessary additional kinetic evidence to further improve the model via Bayesian experimental design. We begin with a presentation of the different computational techniques used here and subsequently discuss the results of our proposed method.

COMPUTATIONAL METHODS

Mean-Field Microkinetic Modeling. A dynamic gas−solid heterogeneous continuous stirred tank reactor (CSTR) microkinetic model was developed to simulate the differential flow reactor and low conversion batch reactor. A high flow rate was maintained, so that the conversion predicted by the model is always less than 1%. The general form of a dynamic CSTR is given below

$$\tau \frac{dF_i}{dt} = F_{in,i} - F_i + \sum_{j \in J} \nu_{ij} r_j \qquad \forall\, i \in I^G \tag{1}$$

$$\frac{d\theta_i}{dt} = \sum_{j \in J} \nu_{ij} r_j \qquad \forall\, i \in I^S \tag{2}$$

$$1 = \sum_{i \in I^S} \theta_i \tag{3}$$

$$r_j = k_{for,j} \Pi_{i \in I_j^{RG}} [P_i]^{\nu_{ij}} \Pi_{i \in I_j^{RS}} [\theta_i]^{\nu_{ij}} - k_{rev,j} \Pi_{i \in I_j^{PG}} [P_i]^{\nu_{ij}} \Pi_{i \in I_j^{PS}} [\theta_i]^{\nu_{ij}} \qquad \forall\, j \in J \tag{4}$$

$$P_i = P_{tot} \frac{F_i}{\sum_G F_l} \qquad \forall\, i \in I^G \tag{5}$$

where i and j are the subscripts over the species set I and reaction set J. F_i, $F_{in,i}$ are the outlet and inlet flow rate of gas species i, and τ is the space time of the CSTR. θ_i is the fractional coverage of surface species i, P_i is the partial pressure of gas species i, and ν_{ij} is the stoichiometric coefficient of species i in reaction j. $k_{for,j}$ and $k_{rev,j}$ are forward and reverse reaction rate constants of a given elementary reaction j. Superscript G and S refer to the gas and surface species, and superscript R and P refer to the reactants and products. P_{tot} is the total reactor pressure. The CSTR model captures the dynamics of the partial pressure of gaseous species (eq 1), the dynamic of the fractional coverage of surface species (eq 2), and the site balance (eq 3). The forward and reverse reaction rate of each elementary reaction follows the law of mass action (eq 4).

The reaction rate constant (k_{for} and k_{rev}) is calculated by the Arrhenius expression

$$k_{for,j} = A_j e^{-E_{aj}/RT} \tag{6}$$

$$k_{for,j} = K_j k_{rev,j} \tag{7}$$

$$K_j = e^{-\Delta H_j/RT} e^{\Delta S_j/R} \tag{8}$$

$$\Delta H_j = \sum_j \nu_{ij} H_i \tag{9}$$

where A_j and E_{aj} is the pre-exponential factor and activation barrier of reaction j, which were calculated by the DFT and transition state (TS) theory. K_j is the equilibrium constant.

Gaussian Process. GP is a supervised non-parametric learning algorithm used for classification and regression, which models the underlying function of the data with a multivariate Gaussian distribution.[18] Gaussian process regression has been used in chemical engineering to calibrate the prediction error of the computational model,[19,20] exploring the reaction networks with uncertainty,[10,21] and modeling the potential energy surface for facilitating the structure optimization and TS search.[22,23] Its non-probabilistic version, namely kernel ridged regression, is used for material screening by predicting the molecular properties[24]

The multivariate Gaussian distribution of GP is defined by the mean and covariance function. The mean function is usually assumed to be zero, and the covariance function estimates the correlation/similarity among the data. Supposing there are n data points, with inputs $\mathbf{X} = (X_1, X_2, ..., X_n)$ and output $\mathbf{y} = (y_1, y_2, ..., y_n)$, given a covariance function $k = k(x,x')$, its prediction y^* on an unseen data point with input X^* follows a joint Gaussian distribution,[18] such as

$$y^*|\mathbf{y} \sim \mathcal{N}(K_* K^{-1}\mathbf{y}, K_{**} - K_* K^{-1} K_*^T) \tag{10}$$

where $\mathcal{N}(\cdot,\cdot)$ stands for multivariate normal distribution giving a mean vector and a covariance matrix. K, K_*, and K_{**} are the kernel matrices, obtained from applying the kernel function on (1) each pair of the inputs ($\mathbf{X}$), (2) the new data and the inputs (X^*, $\mathbf{X}$), and (3) the new data point itself (X^*).

https://dx.doi.org/10.1021/acs.jpcc.0c00491
J. Phys. Chem. C 2020, 124, 5740−5748

The mean of the prediction is a weighted sum of the training data ($K_* K^{-1}\mathbf{y}$), and the weights depend on the similarity among the data points, calculated by the kernel matrix.

The covariance function, also known as the kernel function, is the central part of a GP, which defines the correlation between the data points.[18] The choice of kernel function represents prior knowledge on the underlying function, such as the smoothness, periodicity, and so forth. The squared exponential (SE) kernel, also known as the radial basis function, was chosen as the covariance function here for its smoothness and simplicity in implementation. The SE kernel is expressed as

$$k(x, x') = {\sigma_f}^2 \exp\left(-\frac{\|x - x'\|_2^2}{2l^2}\right) + {\sigma_{\text{noise}}}^2 \tag{11}$$

where l is the characteristic length scale of the input, ${\sigma_f}^2$ and ${\sigma_{\text{noise}}}^2$ are the variance of the kernel and white noise, respectively. Assuming that the likelihood function follows a normal distribution, the analytical expression of the marginalized likelihood is[18]

$$\log P(\mathbf{y}|\mathbf{X}, \theta) = -\frac{1}{2} y^{\mathrm{T}} {K_y}^{-1} y - \frac{1}{2}\log|K| - \frac{n}{2}\log 2\pi \tag{12}$$

where θ is the vector of hyperparameters, calculated via minimizing the negative marginalization likelihood. In this work, we use automatic relevance determination (ARD) SE kernel for each input. The ARD kernel assigns different weights over the features in the input (each dimension in the representation), which automatically determines the relative importance of each feature with hyperparameter optimization.

Bayesian Inference and Experimental Design. Bayesian inference is a statistical framework to update the parameters of a model in accounting for the information from different resources, that is the prior and the evidence.[25] The prior represents the initial belief or guess regarding the parameters, which could come from high level theories, experimental measurements, or expert knowledge. The evidence is a set of new observations of the system. Equation 13 is the mathematical foundation of Bayesian inference, which computes the probability density of the parameters conditioned on the observation.

$$p(\vec{m}|D) = \frac{p(D|\vec{m})p(\vec{m})}{p(D)} \tag{13}$$

$\vec{m}$ is the parameters of the model, and D is the collected data. $p(\vec{m}|D)$ is the posterior distribution of parameters which fuses the information of the prior, $p(\vec{m})$ and evidence, $p(\vec{m})$ through the likelihood function, $p(D|\vec{m})$.

In this study, we use Markov Chain Monte Carlo, a numerical sampling algorithm, to solve this Bayesian inference problem.[25] The Metropolis−Hasting algorithm samples the target distribution by reshaping the proposal distribution to the posterior distribution based on an accept-reject step. In our setting, we run a sufficiently long sampling (more than 100,000 steps) to ensure the convergence of the distribution, which we check by monitoring the mean of the parameters.

The Bayesian experimental design is the probabilistic framework to design/select the most "informative" experiments to refine the model. In this work, we are interested in reducing the prediction uncertainty of the kinetic model, so the experimental data with the largest prediction variance is used to sequentially improve the microkinetic model.

RESULTS AND DISCUSSION

DFT-Based Microkinetic Model Prediction for WGS. The reaction network of the WGS system is shown in Figure 1;

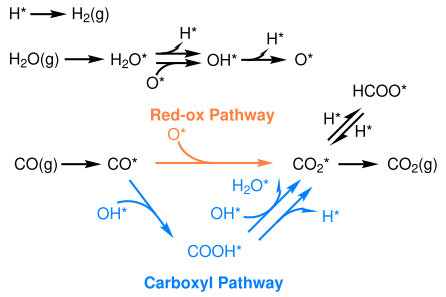


Figure 1. Detailed reaction network of WGS.

the redox pathway (CO* + O* → CO_2* + *) and an associative pathway involving the carboxyl intermediate (via. CO* → COOH* → CO_2*) were included. The temperature-dependent kinetics and thermodynamic parameters are taken from a subset of DFT calculations reported by Grabow and Mavrikakis on Cu(111) using the PW91 functional.[4] Kinetic data reported by Campbell and co-workers on a single-crystal Cu(111) system at 563−683 K, 0.013 atm H_2O, and 0.034 atm CO were used as experimental evidence.[26] Figure 2 shows the comparison of the experimental data with model predictions. The ab initio microkinetic model parameterized with PW91 energies substantially underpredicts the turnover frequencies by 1−4 orders of magnitude, overpredicts the apparent activation barriers by more than 100 kJ/mol, and incorrectly predicts a negative order dependence for CO. Given that (i) both computations and experiments focus on single-crystal facets,[26] (ii) we have covered the most commonly observed pathways reported for the WGS reaction on transition metals, (iii) the coverage of adsorbates predicted by the model was low (≤0.1 ML), so that the mean-field approximation is largely valid, and (iv) most of the surface intermediates (especially, all kinetically relevant ones, i.e., the intermediates and TSs corresponding to steps with a high degree of rate control, see Figures S1 and S2) are strongly bound so that harmonic approximation used by Grabow and Mavrikakis[4] is sufficient to compute their entropies, we argue that this direct comparison between the model and the experiment can be made and the origin of the mismatch largely arises from the inaccuracies in DFT. We note that we implicitly assume that the ZPE energies and entropies calculated by the harmonic oscillator are accurate; in principle, data-driven corrections can also be applied to these terms separately.

GP Calibration on DFT Binding Energies. If indeed the mismatch originates from inaccuracies in DFT, one question that emerges is if these errors of the functional can be learned from related prior thermochemical data, that is not from specific kinetic evidence shown in Figure 2. A more general version of this problem has been considered in the field of Bayesian statistics,[19] viz. if a computer code (or its underlying model/theory) is inadequate to capture the real world, can a correction be applied using related high-fidelity data? One common strategy is to build an external correction around the code (or the model) that is learned from calibration data.

https://dx.doi.org/10.1021/acs.jpcc.0c00491
J. Phys. Chem. C 2020, 124, 5740−5748

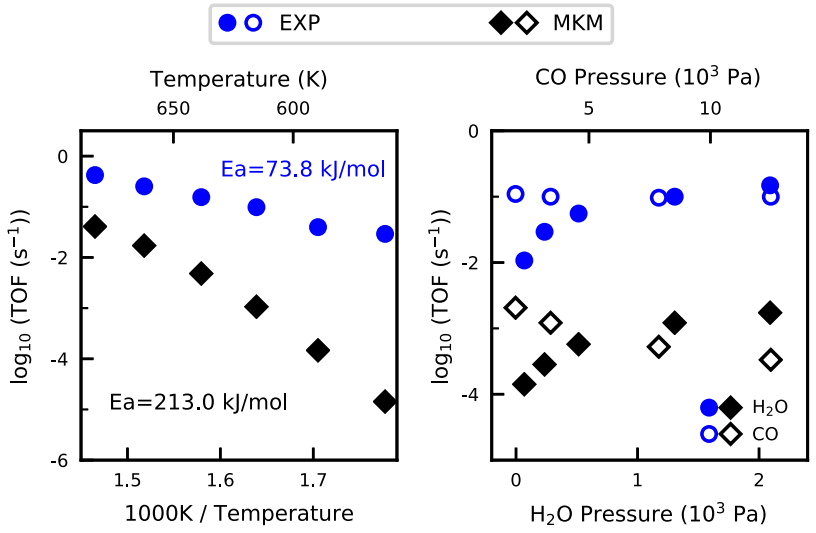


Figure 2. Comparison between experimental[26] and microkinetic model predictions for WGS on Cu(111). (Left) activation barrier is measured at 0.013 atm H_2O and 0.034 atm CO, with temperature from 563 to 683 K; (right) partial pressure of CO is measured at 612 K, 0.013 atm H_2O, varying and the partial pressure of H_2O is measured at 612 K, 0.034 atm CO.

Inspired by this approach, we developed an external calibration model, $f(.)$ to correct the errors in DFT binding energies, $\Delta E_{\mathrm{Exp}}^{\mathrm{DFT}}$ (see eq 14) with experimental data as

$$\Delta E_{\mathrm{Exp}}^{\mathrm{DFT}} = \mathrm{BE}^{\mathrm{DFT}} - \mathrm{BE}^{\mathrm{Exp}} = f(.) \tag{14}$$

where $\mathrm{BE}^{\mathrm{DFT}}$, $\mathrm{BE}^{\mathrm{Exp}}$, and $\Delta E_{\mathrm{Exp}}^{\mathrm{DFT}}$ are the binding energies from DFT calculations and experimental measurements and their difference. Specifically, we used as our data source (i) a subset (~35 points) of the repository of experimental adsorption enthalpies carefully compiled by the Campbell group[6] to get $\mathrm{BE}^{\mathrm{Exp}}$ and (ii) the consistent and homogeneous set of computed binding energies of small adsorbates on a variety of transition metals compiled by the Mavrikakis group using the same software and settings as used earlier (specifically the GGA-PW91 functional) to get $\mathrm{BE}^{\mathrm{DFT=PW91}}$. The dataset is given in the Supporting Information, Section S2; we note here that we subtract computed zero-point energy change values from experimental adsorption enthalpies and treat this as equivalent to the DFT binding energies as a first approximation in view of the low temperatures in many of the experiments.[6]

A GP[18] is trained as an external calibration model ($f(.)$) on $\Delta E_{\mathrm{Exp}}^{\mathrm{PW91}}$. We have previously shown that GP models can be trained to estimate the binding energies of surface adsorbates using simple information about the structure of the adsorbate, the substrate, and the bonding.[27] We use a similar formalism here to predict the error in the DFT estimate of the binding energy given a surface intermediate; furthermore, the GP formalism naturally allows us to calculate a corrected uncertainty estimate of this prediction. The GP installation and the representation of the surface intermediate are shown in the Supporting Information, Section S1.

Figure 3A shows the parity between the GP prediction and the mismatch between the PW91 binding energies and experimental data; the standard deviation of the GP prediction is also given. The model performance is quantified by calculating the root mean squared error (RMSE) and the mean predicted standard deviation (MpSD), which are defined as

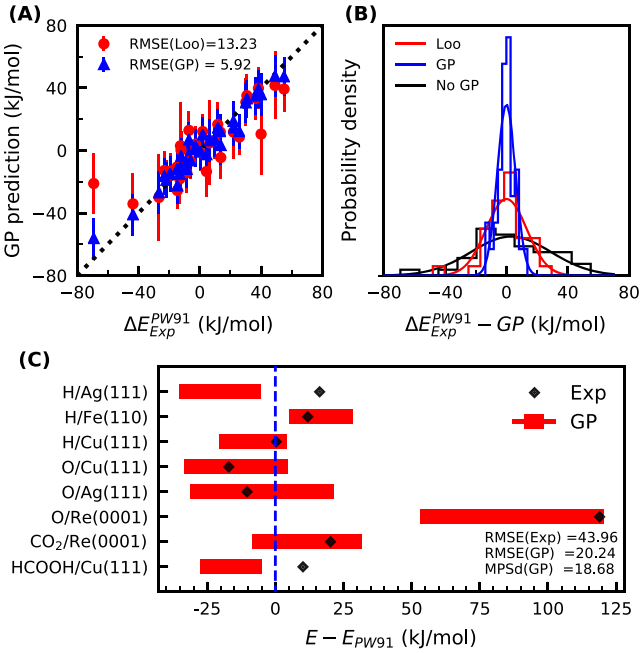


Figure 3. (A) GP prediction vs $\Delta E_{\mathrm{Exp}}^{\mathrm{PW91}}$ (see eq 14 with DFT: PW91). The error bar corresponds to ±1 standard deviation. (B) Error distribution of GP-corrected energies. The histogram is fitted with a normal distribution. (C) GP prediction on the surface species not seen in the training set. The red bar corresponds to ±1 standard deviation. In (A,B), blue corresponds to GP prediction with full dataset as the training set; red corresponds to GP prediction evaluated with leave-one-out cross validation.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_{\mathrm{pred}}^{i} - y_{\mathrm{actual}}^{i})^2} \tag{15}$$

$$\mathrm{MPSd} = \frac{1}{N}\sum_{i=1}^{N}\sigma(y_{\mathrm{pred}}^{i}) \tag{16}$$

where $\sigma(y_{\mathrm{pred}})$ is the predicted standard deviation. Given the paucity of experimental data (35 data points), we train the GP

https://dx.doi.org/10.1021/acs.jpcc.0c00491
J. Phys. Chem. C 2020, 124, 5740−5748

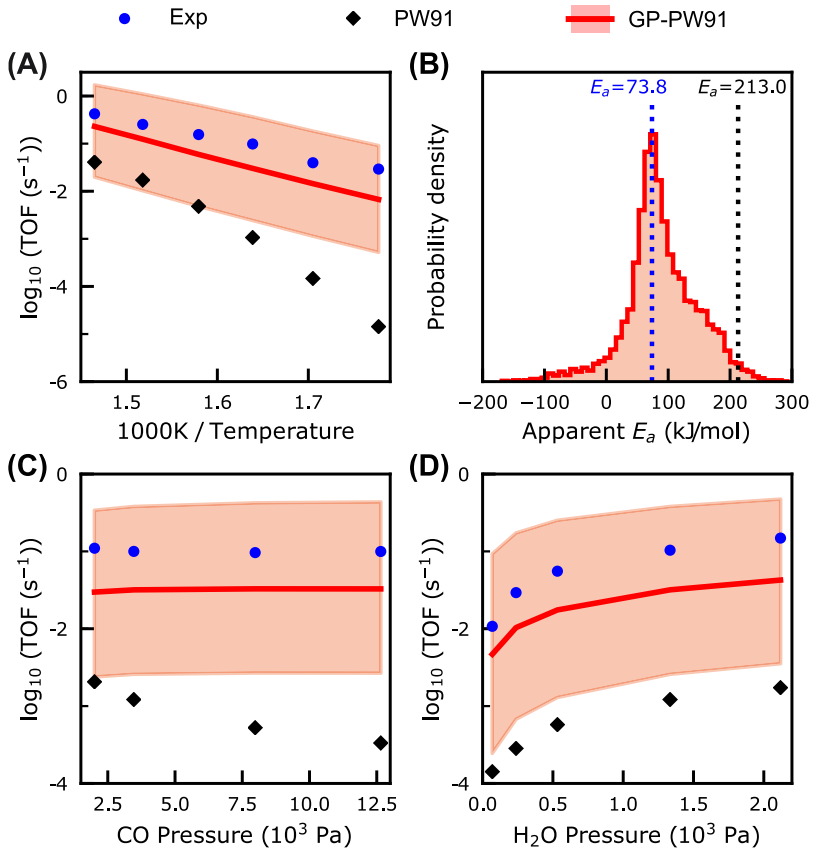


Figure 4. Comparison between experimental data[26] and kinetic models for water-gas shift on Cu(111). (A) Arrhenius plot; (B) distribution of apparent activation energy; (C) TOF vs partial pressure of CO; and (D) TOF vs partial pressure of H_2O. Blue: kinetic experiments; black: kinetic model prediction with original PW91 energies; red: kinetic model prediction with GP-corrected PW91 energies, the red line: median (50%) of the prediction, the shaded region corresponds to 16−84 percentile prediction.

on all data points, and the RMSE is 5.9 kJ/mol. We performed a leave-one-out cross validation (LOO-CV) to evaluate the generalization error. The root mean square error of this exercise (LOO-RMSE) is 13 kJ/mol, which is higher than the previous case, indicating that GP overfits this limited dataset. Nevertheless, as long as we do not extrapolate far beyond the current chemical space, the GP prediction should be reliable. As shown in Figure 3B, the distribution of the errors of the GP-corrected energies is narrower than that of PW91 (whose standard deviation is 26 kJ/mol); GP-corrected energies are therefore more reliable. In Figure 3C, to verify the performance of the GP, we compared the GP prediction with a set of additional experimental data (outside the training set), and it shows that GP can capture the trends in the intrinsic error in DFT estimates, and the predicted variance provides a good estimation of the model error (see Table S2 for the testing data). Furthermore, Figure S3 in the Supporting Information, which shows a heat map of the covariance matrix obtained from the GP model, reveals physically meaningful similarities between adsorbate−surface combinations. For instance, the model correctly identifies that the binding of OH* and CH_3O* on Pt(111) is similar relative to, say, OH* and bidentate HCOO*. We note here that there are potential uncertainties associated with the coverage of adsorbates in experimental single-crystal studies and that we have not explicitly taken this into account in our model; however, as the coverages in both the experiments and the reference DFT calculations are low (<0.25 ML) for the cases considered here, we expect them to not play a significant role. Despite the caveats, the qualitative and quantitative performance of the GP corrections to original DFT predictions allow us to employ this model with confidence in subsequent analyses.

Applying the GP Corrections to the Microkinetic Model. We checked the performance of the GP corrections on the mean-field microkinetic model for the WGS reaction. The GP model corrects the DFT binding energies and provides an uncertainty distribution associated with it (Supporting Information, Section S3.1). The corrections for TS energies of each elementary reaction were estimated by the proximity method adopted by Grabow and Mavrikakis,[4,28] where the deviation of the TS energy from DFT is taken to be linearly dependent on the deviation of initial and final state energies (details of the implementation are given in the Supporting Information, Section S1.3). Upon correcting the DFT energies with the GP model, shown in red in Figure 4A−D, the resulting model can, a priori, capture the experimental trends (turnover frequency, apparent barriers, and reaction orders) significantly more accurately than the original model. Furthermore, the distribution of the GP was sampled and propagated through the microkinetic model to get an uncertainty band on the predictions. As shown in the figure, the predicted distribution of the GP-corrected microkinetic model spans 2 orders of magnitude because the predicted variance of GP calibration model is around 10−20 kJ/mol. Clearly, augmenting the DFT energies with GP corrections largely "fixes" the mismatch between experiments and the original model, thereby lending credence to our argument that

https://dx.doi.org/10.1021/acs.jpcc.0c00491
J. Phys. Chem. C 2020, 124, 5740−5748

the mismatch arises from errors in the PW91 functional in this case.

Application of the Method to Other Catalytic Systems. We also employed GP-corrected microkinetic modeling for other catalytic systems on single-crystal Cu(111) surface, specifically, TPD of $H_2(g)$ and $H_2O(g)$, methanol synthesis, and reverse WGS reactions. In methanol synthesis and reverse WGS reaction, the microkinetic model parameterized with PW91 energies underestimates the TOF of $CH_3OH(g)$ and $CO(g)$ by 6−8 orders of magnitude and overestimates the selectivity of $CH_3OH(g)$ over $CO(g)$ ($S_{CO}^{CH_3OH} > 1$). After augmenting GP correction, the predictions improve by 1−3 orders of magnitude, although still underpredicting the experimentally observed rates. The prediction of the selectivity is improved ($S_{CO}^{CH_3OH} < 0.01$). This result suggests that GP captures the relative errors between the selectivity determining steps and underestimates the error in the overall rate limiting step. For the TPD of $H_2(g)$ and $H_2O(g)$, we focus on the peak position of the temperature, which is a quantity directly related to the binding energy of the surface intermediate (BE^H and BE^{H_2O}). In TPD of $H_2(g)$, the microkinetic model with PW91 energies underestimates the peak temperature by around 100 K, whereas post corrections, the predictions improved by 50 K. For the TPD of $H_2O(g)$, the microkinetic model initially underpredicts by over 120 K; with GP correction, the model prediction was within 20 K of the experiments.

All three cases show improvements compared to DFT-based microkinetic model predictions; the extent of improvements, we can note, varies from one system to another and depends on the extent to which the GP corrections for kinetically relevant intermediates and TSs were accurate. The readers are referred to the Supporting Information, Section S3.3 for more details.

Posterior Inference with Experimental Evidence. We note that our GP-corrected microkinetic model is an a priori kinetic model in that no kinetic evidence pertaining to the WGS reaction has been used yet. Nevertheless, the correlated uncertainty distributions offered by GP could be used as an informed prior to identify most informative kinetic experiments that can be used as evidence to further refine the model, to improve its predictive power, and to shrink its uncertainty bands. Once refined with the appropriate kinetic data, the model is postdictive, but can be substantially more robust. We use Bayesian inference and design of experiments to this end, which allows us to incorporate information from new data sources to update the existing knowledge about a reaction system. Previously, we have shown that the parameters of a kinetic model can be re-estimated using rigorous nonlinear optimization.[7] The solution of such problems is a point estimate of the parameters in the solution space (corresponding to a local minimum of a least squares objective), although simple multistart strategies can be deployed in an embarrassingly parallel mode to obtain a slew of alternative solutions. Bayesian inference, on the other hand, provides a correlated posterior probability distribution of the parameters.[25] The distribution then reflects an intrinsic uncertainty in the parameter value and every sample represents an alternative solution with a finite probability. The advantage of this procedure is that one can then use "posterior" distribution of the parameters and the model prediction (i.e., after modifying the prior knowledge of the parameters with new evidence) to identify where to collect more evidence (i.e., new experimental data). One approach is to consider a collection of potential experimental conditions and evaluate where the model has the largest prediction variance, pick that condition for the next experiment, and then feed this new evidence to retrain the model. In our illustration here, as the data are already available and this modeling exercise is conducted after the fact, we demonstrate how this approach could generally work in real time to further refine a microkinetic model.

Figure 5 demonstrates that as we refine the kinetic model using Bayesian experimental design, the model quickly

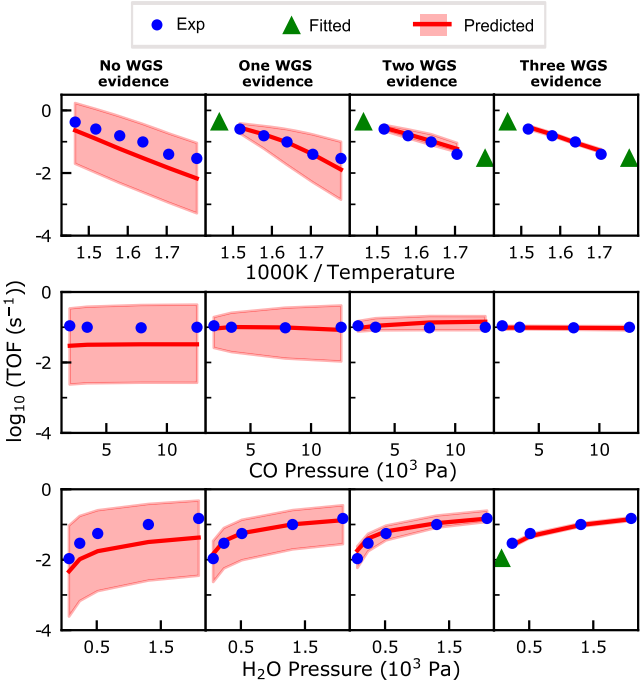


Figure 5. Bayesian experimental design of WGS kinetic with GP as prior. First row: varying temperature. Second row: varying partial pressure of CO(g), third row: varying partial pressure of $H_2O(g)$. Blue dot: unseen (unused) experimental data. Green triangle: fitted data. Red shade: prediction on unseen data.

converges. Specifically, the microkinetic model reproduces all experimental data with only two/three experimental data points. Additionally, the systematic reduction in the prediction variance indicates that the posterior distribution captures the information from the prior (GP calibration model) and new evidence (WGS kinetic data). There is a clear trend that the predicted variance of all unseen (or unused) kinetic data is progressively decreasing as well. In addition, the decrease in the predicted variances of unknown reaction conditions is not uniform; the uncertainty bands are larger at points away from the fitted data.

Figure 6A compares the potential energy surfaces (PES) corresponding to three versions of the microkinetic models presented thus far—one based on original DFT energies, one after GP corrections, and the third being the posterior model. Augmenting the DFT energies with the GP model (GP-PW91, red) shifts the PES downward by binding all surface intermediates more strongly. After inferring from three kinetic data points, the distribution of the PES (green) marginally shifts from that of GP-PW91, and the variance of the PES decreases. Figure 6B shows the net reaction rates of two

https://dx.doi.org/10.1021/acs.jpcc.0c00491
J. Phys. Chem. C 2020, 124, 5740−5748

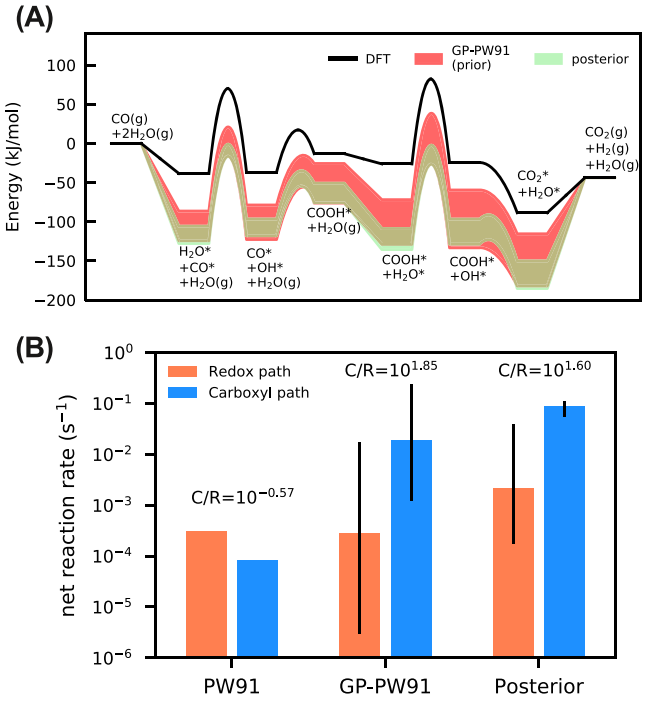


Figure 6. (A) Potential energy surface of WGS on copper (111). Black: origin PW91. Red shade: GP-PW91. Green shade: posterior with three kinetic experimental data. The uncertainty corresponds to ±1 standard deviation. H* is omitted from labels. (B) Comparison of the net reaction rate of two pathways predicted by PW91, GP augmented PW91, and posterior distribution with three kinetic evidences. C/R stands for the ratio of the reaction fluxes of carboxyl and redox pathway.

reaction pathways predicted by these three microkinetic model versions. The microkinetic model with PW91 energies initially predicts that the redox pathway is dominant. After including the GP-based corrections, the resulting model predicts that the carboxyl pathway is the dominant one, which is consistent with previous predictions;[12] further incorporation of kinetic evidence does not alter this. The predicted variance of two pathways decreases after inferring from kinetic data. In addition, the variance of the predicted flux through the carboxyl pathway in the posterior model is a lot smaller than that of the redox pathway indicating that some elementary reactions (in the main pathway) are tightly connected to the overall kinetics (TOF) than others. The uncertainty corresponding to those reactions reduce upon including more kinetic data. However, the reaction fluxes of side reactions are not directly inferred from the overall TOFs, so the uncertainties within those reactions are not significantly reduced. The detailed predictions of reaction fluxes, the value of the degree of rate control of each step, and surface coverage are given in the Supporting Information, Section S3.1. The microkinetic model based on PW91 energies predicted a large positive degree of rate control (~2.0) for OH* dissociation and CO* oxidation (of the redox pathway), whereas the GP-PW91 and the final posterior model predict OH* dissociation to be rate determining. All models indicate that the surface is largely clean under reaction conditions.

Figures 7 and 8 compare the distribution of the activation barriers of reactions and binding energies of intermediates across different models. The predicted standard deviation of the GP is wide. Upon inferring from kinetic experimental data, the width of the distribution shrinks. Comparing GP-PW91 with the posterior model (inferred from three kinetic data), we can note that the individual distributions shift by 0−15 kJ/mol. The two posterior distributions, inferred with three and fifteen experimental data values, are quite close in most cases (except for the reaction COOH* + OH* → CO_2* + H_2O*). These two observations further confirm that the GP-corrected model had sufficiently correctly captured the chemistry and that three additional kinetic data points are sufficient to refine the model to an acceptable level of accuracy. We can therefore verify our original supposition that the chemistry and active sites were correctly captured in our model and the original model-experiment mismatch shown in Figure 2 is entirely due to intrinsic errors in the chosen functional.

CONCLUSIONS

In summary, we present a method to systematically improve microkinetic models of catalytic systems. We train a GP model as external correction to DFT binding energies using, albeit sparingly, available experimental single-crystal adsorption

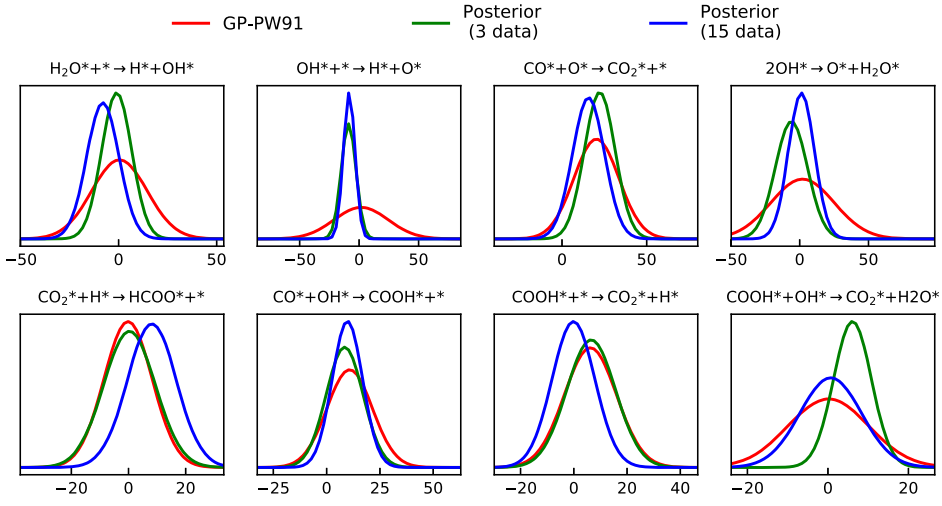


Figure 7. Distribution of the deviation on E_a of all elementary reactions in WGS chemistry (units in kJ/mol). Red: prior distribution from GP; green: posterior distribution inferred with three kinetic data; Blue: posterior distribution inferred with fifteen kinetic data.

https://dx.doi.org/10.1021/acs.jpcc.0c00491
J. Phys. Chem. C 2020, 124, 5740−5748

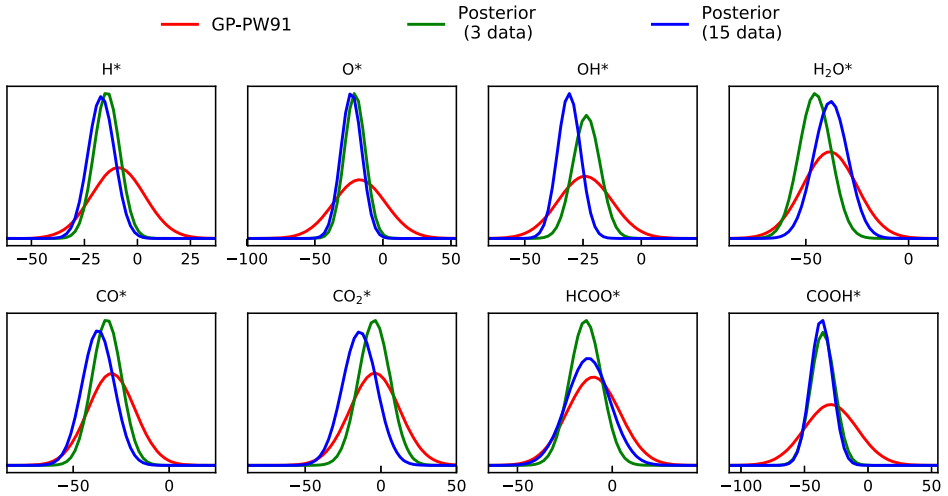


Figure 8. Distribution of the deviation on binding energies of all surface intermediates in WGS chemistry (units in kJ/mol). Red: prior distribution from GP; green: posterior distribution inferred with three kinetic data; blue: posterior distribution inferred with fifteen kinetic data.

enthalpies. We show that the inclusion of this statistical correction to the kinetic and thermodynamic parameters of an illustrative microkinetic model of the WGS chemistry on Cu(111) (and many more) substantially improves prediction accuracy. The correlated uncertainties resulting from the GP model can be propagated forward to obtain the variance of the prediction of the microkinetic model, and also serve as an informative prior for subsequent Bayesian experimental design to iteratively refine the model. A pertinent challenge in adopting this method is that adequate and relevant experimental thermochemical data may not be available for many classes of catalytic materials or chemistries to train a GP model on. We offer the concept of multifidelity modeling as a partial remedy to this problem. Specifically, we have shown that a data-driven binding energy model can be simultaneously trained on "large" amounts of (relatively) low-fidelity (such as GGA functionals) and relatively small amount of high-fidelity data (such as the random phase approximation).[27] The resulting model has a higher accuracy than a model trained on either sets of data alone. One could then build a corrective model to DFT by training a GP model on the difference between high- and low-fidelity energies or build a multifidelity model to estimate the binding energies directly; the resulting model could then be used to parameterize a microkinetic model. Nevertheless, for metal catalysis chemistries involving small to medium intermediates, for which there is sufficient thermochemistry data,[29] this is a simple and viable strategy to correct microkinetic models in the absence of any kinetic evidence.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcc.0c00491.

Computational setting of GP, representation of surface intermediates, and proximity method for estimation activation barrier, and prediction of microkinetic modeling; benchmarking and testing dataset for GP, kinetic experimental data; additional results about GP prediction, microkinetic prediction of WGS (forward simulation and Bayesian inference), and the application of GP on other catalytic systems are available in the supporting information. The example code about GP training, forward uncertainty propagation, and Bayesian inference of the microkinetic model for WGS reaction on Cu(111) are available in WGS_GP repository at https://github.com/thj2009/WGS_Cu (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Srinivas Rangarajan − *Department of Chemical and Biomolecular Engineering, Lehigh University, Bethlehem, Pennsylvania 18015, United States;* orcid.org/0000-0002-6777-9421; Phone: +1 610-758-4219; Email: srr516@lehigh.edu

Author

Huijie Tian − *Department of Chemical and Biomolecular Engineering, Lehigh University, Bethlehem, Pennsylvania 18015, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcc.0c00491

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was partially supported by startup funds at Lehigh University. S.R. further acknowledges partial support from the American Chemical Society Petroleum Research Fund. Grant number: #59476 DNI5.

■ REFERENCES

(1) Dumesic, J. A. *The Microkinetics of Heterogeneous Catalysis*; American Chemical Society: Washington, DC, 1993.

(2) Gokhale, A. A.; Kandoi, S.; Greeley, J. P.; Mavrikakis, M.; Dumesic, J. A. Molecular-level descriptions of surface chemistry in kinetic models using density functional theory. *Chem. Eng. Sci.* **2004**, *59*, 4679−4691.

(3) Nørskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. Towards the computational design of solid catalysts. *Nat. Chem.* **2009**, *1*, 37−46.

(4) Grabow, L. C.; Mavrikakis, M. Mechanism of methanol synthesis on Cu through CO2 and CO hydrogenation. *ACS Catal.* **2011**, *1*, 365−384.

(5) Matera, S.; Schneider, W. F.; Heyden, A.; Savara, A. Progress in accurate chemical kinetic modeling, simulations, and parameter estimation for heterogeneous Catalysis. *ACS Catal.* **2019**, *9*, 6624−6647.
(6) Wellendorff, J.; Silbaugh, T. L.; Garcia-Pintos, D.; Nørskov, J. K.; Bligaard, T.; Studt, F.; Campbell, C. T. A benchmark database for adsorption bond energies to transition metal surfaces and comparison to selected DFT functionals. *Surf. Sci.* **2015**, *640*, 36−44.
(7) Rangarajan, S.; Maravelias, C. T.; Mavrikakis, M. Sequential-optimization-based framework for robust modeling and design of heterogeneous catalytic systems. *J. Phys. Chem. C* **2017**, *121*, 25847−25863.
(8) Tameh, M. S.; Dearden, A. K.; Huang, C. Accuracy of density functional theory for predicting kinetics of methanol synthesis from CO and CO2 hydrogenation on copper. *J. Phys. Chem. C* **2018**, *122*, 17942−17953.
(9) Bukowski, B. C.; Bates, J. S.; Gounder, R.; Greeley, J. First principles, microkinetic, and experimental analysis of Lewis acid site speciation during ethanol dehydration on Sn-Beta zeolites. *J. Catal.* **2018**, *365*, 261−276.
(10) 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.
(11) Salciccioli, M.; Stamatakis, M.; Caratzoulas, S.; Vlachos, D. G. A review of multiscale modeling of metal-catalyzed reactions: Mechanism development for complexity and emergent behavior. *Chem. Eng. Sci.* **2011**, *66*, 4319−4355.
(12) Gokhale, A. A.; Dumesic, J. A.; Mavrikakis, M. On the mechanism of low-temperature water gas shift reaction on copper. *J. Am. Chem. Soc.* **2008**, *130*, 1402−1414.
(13) Sutton, J. E.; Guo, W.; Katsoulakis, M. A.; Vlachos, D. G. Effects of correlated parameters and uncertainty in electronic-structure-based chemical kinetic modelling. *Nat. Chem.* **2016**, *8*, 331−337.
(14) Medford, A. J.; Wellendorff, J.; Vojvodic, A.; Studt, F.; Abild-Pedersen, F.; Jacobsen, K. W.; Bligaard, T.; Norskov, J. K. Assessing the reliability of calculated catalytic ammonia synthesis rates. *Science* **2014**, *345*, 197−200.
(15) Walker, E.; Ammal, S. C.; Terejanu, G. A.; Heyden, A. Uncertainty quantification framework applied to the watergas shift reaction over Pt-based catalysts. *J. Phys. Chem. C* **2016**, *120*, 10328−10339.
(16) Wellendorff, J.; Lundgaard, K. T.; Mgelhj, 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.
(17) Walker, E. A.; Mitchell, D.; Terejanu, G. A.; Heyden, A. Identifying active sites of the watergas shift reaction over titania supported platinum catalysts under uncertainty. *ACS Catal.* **2018**, *8*, 3990−3998.
(18) Rasmussen, C. E.; Williams, C. K. I. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*; The MIT Press, 2005.
(19) Kennedy, M. C.; O'Hagan, A. Bayesian calibration of computer models. *J. Roy. Stat. Soc. B Stat. Methodol.* **2001**, *63*, 425−464.
(20) Pernot, P.; Cailliez, F. A critical review of statistical calibration/prediction models handling data inconsistency and model inadequacy. *AIChE J.* **2017**, *63*, 4642−4665.
(21) Simm, G. N.; Reiher, M. Error-controlled exploration of chemical reaction networks with Gaussian processes. *J. Chem. Theory Comput.* **2018**, *14*, 5238−5248.
(22) Schmitz, G.; Christiansen, O. Gaussian process regression to accelerate geometry optimizations relying on numerical differentiation. *J. Chem. Phys.* **2018**, *148*, 241704.
(23) Garrido Torres, J. A.; Jennings, P. C.; Hansen, M. H.; Boes, J. R.; Bligaard, T. Low-Scaling algorithm for nudged elastic band calculations using a surrogate machine learning model. *Phys. Rev. Lett.* **2019**, *122*, 156001.
(24) Rupp, M.; Tkatchenko, A.; Mller, K.-R.; von Lilienfeld, O. A. Fast and accurate modeling of molecular atomization energies with machine learning. *Phys. Rev. Lett.* **2012**, *108*, 058301.
(25) Gelman, A.; Carlin, J. B.; Stern, H. S.; Dunson, D. B.; Vehtari, A.; Rubin, D. B. *Bayesian Data Analysis*; Chapman and Hall/CRC, 2013.
(26) Campbell, C.; Daube, K. A. A surface science investigation of the water-gas shift reaction on Cu(111). *J. Catal.* **1987**, *104*, 109−119.
(27) Tian, H.; Rangarajan, S. Predicting adsorption energies using multifidelity data. *J. Chem. Theory Comput.* **2019**, *15*, 5588−5600.
(28) Grabow, L. C.; Gokhale, A. A.; Evans, S. T.; Dumesic, J. A.; Mavrikakis, M. Mechanism of the water gas shift reaction on Pt: First principles, experiments, and microkinetic modeling. *J. Phys. Chem. C* **2008**, *112*, 4608−4617.
(29) Campbell, C. T. Energies of adsorbed catalytic intermediates on transition metal surfaces: Calorimetric measurements and benchmarks for theory. *Acc. Chem. Res.* **2019**, *52*, 984−993.

https://dx.doi.org/10.1021/acs.jpcc.0c00491
J. Phys. Chem. C 2020, 124, 5740−5748