



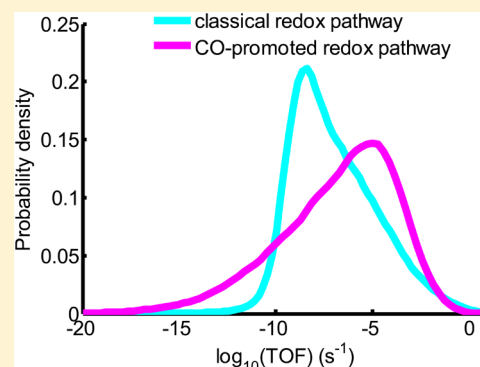
Uncertainty Quantification Framework Applied to the Water–Gas Shift Reaction over Pt-Based Catalysts

Eric Walker,[†] Salai Cheettu Ammal,[†] Gabriel A. Terejanu,^{*,‡} and Andreas Heyden^{*,†}

[†]Department of Chemical Engineering, University of South Carolina, 301 Main Street, Columbia, South Carolina 29208, United States

[‡]Department of Computer Science and Engineering, University of South Carolina, 301 Main Street, Columbia, South Carolina 29208, United States

ABSTRACT: This paper presents a systematic approach to quantify uncertainties of various quantities of interest (QoIs) in catalysis determined by microkinetic models developed from first principles. One of the main sources of uncertainty in any microkinetic simulation is attributed to the exchange-correlation approximations in density functional theory (DFT) used to calculate the rate constants for all elementary reaction steps within transition state theory. These DFT approximations are at the core of significant discrepancies between computational simulations and experimental measurements. Therefore, any model calculation should be accompanied by a measure of uncertainty. This work uses probability to represent uncertainties and latent variable models to develop probabilistic models that account for errors and correlations in DFT energies. These probabilistic models are further constrained to known reaction thermodynamics, and then propagated to QoIs such as turnover frequency (TOF), apparent activation barrier, and reaction orders. The proposed uncertainty quantification (UQ) framework is applied on the water–gas shift reaction (WGS: $\text{CO} + \text{H}_2\text{O} \rightleftharpoons \text{CO}_2 + \text{H}_2$). Specifically, this WGS study models a Pt_8 cluster supported on a rutile TiO_2 (110) surface, where DFT energies are obtained using four separate functionals PBE, RPBE, HSE, and M06L that each have their own justification for being appropriate for this study. In this way, information from three different classes of functionals, GGA (generalized-gradient approximation), meta-GGA, and hybrid functionals, are used to generate a free energy probabilistic model. Although the uncertainty in model results spans orders of magnitude, a new approach is introduced to identify the dominant catalytic cycle under uncertainty. Overall, we find that our model captures various experimental kinetic data; however, the probability densities for TOF, apparent activation barrier, and reaction orders are relatively wide due to different flavors of DFT predicting a wide variation of transition state and oxygen vacancy formation energies. Nevertheless, we can conclude with high certainty that a CO-promoted redox cycle is the dominant mechanism over the temperature range 473–600 K and that formate and carboxyl pathways are not playing any role for the investigated active site model.



1. INTRODUCTION

Microkinetic models play an important role in understanding reaction kinetics and production rates on macroscopic scales. The insight provided by microkinetic model simulations can be used to speed up the rational design of novel catalytic materials.^{1–3} Density functional theory (DFT) provides a good starting point to estimate (together with harmonic transition state theory) the values of rate constants for each elementary step in the microkinetic model. However, DFT calculations have inadequacies due to exchange-correlation approximations, which induce significant errors in the prediction of macroscopic quantities of interest (QoI) such as turnover frequency (TOF), apparent activation barrier, and reaction orders. For example, a variation of only 0.2 eV in a DFT predicted activation barrier leads to an uncertainty of 2 orders of magnitude in an elementary reaction rate constant in a catalytic cycle at 500 K, which again can lead to significant uncertainty in microkinetic modeling results. Therefore, new computational tools are

required to quantify uncertainties in DFT calculations, propagate them to QoIs, and guide the process of drawing conclusions under uncertainty. We note that we limit ourselves in the following to uncertainties in the DFT functional and neglect any uncertainties related to estimating entropies and using harmonic transition state theory. In fact, we assume that these uncertainties are small relative to the uncertainties in DFT energies such that they can approximately be mapped into the uncertainties of the DFT energies (an approximation that might not be valid for some surface states). We note that considering these uncertainties does not constitute any additional complication and our proposed framework can in principle be extended to include these uncertainties. Also, we are not concerned with inadequacies of our active site model structure

Received: February 8, 2016

Revised: April 26, 2016

Published: April 27, 2016

and number of elementary reaction steps but aim at predicting the QoIs for a given active site model and reaction mechanism.

Previously, a number of research groups such as Mortensen et al.,⁴ Cramer,⁵ and Hanke⁶ have investigated the errors in DFT energies of functionals commonly used in heterogeneous catalysis. Also, Vlachos and co-workers^{7,8} have proposed corrections for energies of formation and binding energies obtained by DFT calculations to explain differences in DFT energies and experimental observations. Errors in DFT energies have often been assumed to behave systematically; e.g., a constant correction term was assigned to any species containing an OCO backbone.⁹ This systematic error was determined by comparing DFT energies of gas molecules to National Institute of Standards and Technology (NIST) data.¹⁰ While such a systematic approach is appealing, there is little scientific justification regarding why errors in DFT energies can be divided into atom groups that are transferable to systems containing transition metal surfaces. Recently, Petzold et al.¹¹ have proposed the use of databases of electronic density functionals fit by Bayes formula. This research led to the Bayesian error estimation–van der Waals (BEEF-vdW) functional,^{4,12–14} a seminal work in quantifying uncertainties in DFT energies. BEEF-vdW is based on using empirical data of microscopic properties such as binding energies to infer errors and correlations in DFT energies due to the exchange–correlation functional approximation. Medford et al.¹² recently found by studying the ammonia synthesis reaction that these correlations in energies play an important role in reducing the uncertainty in QoIs such as TOFs. One way to obtain correlations is to use the ensemble or a subset of the ensemble of functionals from BEEF-vdW.^{4,12–14} However, these correlations are based strictly only on generalized-gradient approximation (GGA) functionals, and GGAs can face issues when electrons become localized such as in oxide systems.^{15,16} For example, in the system of consideration, oxygen vacancies are formed on a TiO₂ surface, and a strategy based exclusively on GGA functionals for estimating errors and their correlation risks underestimating the inherent uncertainties since all GGA functionals will overdelocalize electrons which leads to lower vacancy formation energies.

In this work, we propose to describe the uncertainty in DFT energies by explicitly accounting for information provided by various classes of functionals such as GGA, meta-GGA, as well as hybrid functionals. Specifically, we only choose DFT functionals (PBE,¹⁷ RPBE,^{18,19} HSE,²⁰ and M06L²¹) that have some justification to be used in a DFT study of the active site model structure and that are all known to have different limitations relevant for the reaction mechanism. Next, DFT energies calculated with these functionals are used in a factor analysis²² to develop a probabilistic latent variable model that accounts for errors and correlations in DFT energies.

This probabilistic model is further extended to ensure that samples of DFT energies are constrained to known reaction thermodynamics. This is accomplished by designing a separate probabilistic model for energy correction of gas molecules based on the Dirichlet distribution.²³ The uncertainty captured by this composite probabilistic model for DFT energies is propagated through the microkinetic model to QoIs using Monte Carlo simulations. Specifically, the composite probabilistic model is used to generate correlated and thermodynamically constrained DFT energy samples, which are then used to calculate rate constants in the microkinetic model. Finally, the uncertainty in the QoIs such as TOF, reaction orders, and

apparent activation energy is captured by samples corresponding to the QoI predictive distribution. As the model calculations are no longer deterministic but rather probabilistic, new tools are required to draw conclusions under uncertainty. This work introduces the information theoretic quantity, Kullback–Leibler divergence, to determine the dominant pathway by finding the smallest divergence between the probability density function (PDF) of the overall TOF and the PDF of individual pathway TOF.

To summarize, uncertainty quantification (UQ) is the process of assessing and representing uncertainties in model simulations such that their impacts on the QoIs can be determined. This paper presents a systematic approach to quantify uncertainties in QoIs calculated using microkinetic models. The proposed UQ framework for computational catalysis consists of three distinct processes: (1) quantifying uncertainties in DFT energies, (2) propagating uncertainties to QoIs, and (3) drawing conclusions using uncertain QoIs. To showcase the proposed UQ framework, we apply it to our recent water–gas shift reaction model (WGS: CO + H₂O ⇌ CO₂ + H₂) over Pt/TiO₂ catalysts.²⁴

The WGS is an essential step in industrial chemical processes utilizing hydrogen.^{24–33} Besides producing hydrogen, the WGS consumes carbon monoxide (CO) which is beneficial because CO is a poison for noble metal catalysts and catalysts used in fuel cell applications.^{26,32,33} With these important applications comes a desire to understand the WGS over relevant catalyst systems. The case study in this work builds upon a theoretical investigation²⁴ of the WGS over an active site at a three-phase boundary (TPB) of a metal nanoparticle, a reducible metal oxide support, and a gas phase. It has been hypothesized that TPB sites are the origin of unique activity and selectivity of various reducible oxide supported metal cluster catalysts. For example, gold (Au) catalysts on reducible oxide supports possess a unique activity attributed to TPB sites.^{34–36} The specific catalyst of this case study is platinum supported by titanium oxide (Pt/TiO₂).^{37–39} Pt/TiO₂ reports more activity than systems of Pt on other mixed oxides³⁴ and has been investigated for low and medium temperature WGS reactions.^{40–47} A current question of high interest is if individual metal atoms or metal clusters/nanoparticles supported on reducible oxide supports are the active site for catalysis.^{48–50} DFT calculations and microkinetic modeling can be used to shed more light on the activity of various active site models and has been used here for a Pt₈/TiO₂ (110) catalyst model characteristic of a metal cluster on rutile TiO₂. At the same time, describing the electronic structure of reaction intermediates and transition states at the interface of a Pt cluster and a TiO₂ support during the WGS is quite challenging with DFT considering that regions of the active site model are clearly metallic while the oxide support is a semiconductor. Considering furthermore that the oxygen vacancy formation energy of TiO₂ is significantly underpredicted by GGA-DFT,⁵¹ there is increased uncertainty in our previous computational predictions that have been based on GGA-DFT,^{24,52} making this catalytic system an ideal test case for our novel UQ approach.

2. UNCERTAINTY QUANTIFICATION

This section introduces the proposed UQ framework for a general system where *N* functionals have been used to obtain energies for intermediate and transition states. In this work, *N* = 4. This data set is used to generate a probabilistic latent variable model to capture both the variation of DFT energies

for individual states and the correlation between energies at various states. This probabilistic model is further constrained to known reaction thermodynamics, and then it is propagated to QoIs such as turnover frequency (TOF), apparent activation barrier, and reaction orders. Finally, an information theoretic approach is introduced to determine the dominant catalytic cycle under uncertainty.

2.1. DFT Latent Variable Model. In this work the probabilistic latent variable model used to summarize the uncertainty in DFT calculations is developed using factor analysis. Factor analysis expresses observed variables (functional results) as a linear combination of a small number of factors.²² This allows the model to represent the uncertainty in the functionals as a multivariate Gaussian with a restricted number of free parameters and still capture the dominant correlations in the data set. Using fewer factors than functionals is possible because the functionals are correlated.^{4,9} An exchange-correlation functional in DFT tends to overestimate one intermediate state energy if it overestimates a similar intermediate energy. Positive correlation signifies corrections moving together, and negative correlation corresponds to corrections occurring opposite to each other.

Factor analysis models the observable in the following manner.

$$\mathbf{y} = \mathbf{W}\mathbf{z} + \boldsymbol{\mu} + \mathbf{e} \quad (1)$$

Here, the vector \mathbf{y} denotes the observed DFT Gibbs' free energies for all intermediate and transition states relative to a clean catalyst with reactant molecules in the gas phase, and \mathbf{z} is a vector of independent factors/latent variables that are normally distributed, $N(0, 1)$. The matrix \mathbf{W} contains the factor loadings that capture the correlations between observables, and \mathbf{e} is a zero-mean Gaussian distributed noise with a diagonal covariance matrix $\boldsymbol{\Psi}$ that contains *specific variances*.²²

The parameters, \mathbf{W} , $\boldsymbol{\mu}$, and $\boldsymbol{\Psi}$ are determined using the maximum likelihood²² via expectation maximization; i.e., $\boldsymbol{\mu}$ becomes a vector of the average of the DFT predicted Gibbs' free energies.⁵³ Given the set of N DFT calculations, $\mathbf{Y} = \{y_1, \dots, y_N\}$, the corresponding log-likelihood function, $\ln p$, to be maximized is given by

$$\ln p(\mathbf{Y}|\mathbf{W}, \boldsymbol{\mu}, \boldsymbol{\Psi}) = \sum_{i=1}^N \ln p(y_i|\mathbf{W}, \boldsymbol{\mu}, \boldsymbol{\Psi}) \quad (2)$$

The number of optimal factors is determined using cross-validation. After parameter learning, the marginal distribution for the observable y is given by a Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix given by $\boldsymbol{\Sigma} = \mathbf{W}\mathbf{W}^T + \boldsymbol{\Psi}$. With this method the correlation structure is from N functional results rather than setting a correlation structure by choosing a chemical similarity measure among states and transition states. This latent variable model is used to draw samples of energies corresponding to intermediate and transition states. However, these energy samples may not match the reaction thermodynamics.

2.2. Thermodynamics Correction. The factor analysis described above has been conducted on the free energies of surface intermediate and transition states. After the factor analysis is completed, a separate probabilistic model is proposed to correct the free energies of the gas molecules such that they match the overall reaction thermodynamics (Gibbs' free energy of reaction). Given M gas molecules with $\{\zeta_1, \zeta_2, \dots, \zeta_M\}$

corrections, then the following constraint is imposed on the gas molecule corrections.

$$\sum_{i=1}^M \zeta_i = \delta \quad (3)$$

Note that δ is constant (the average error of our DFT functionals in predicting the Gibbs' free energy of reaction). The individual corrections may be bounded due to the uncertainty associated with the magnitude of the correction. Namely, each correction has its own associated uncertainty.

$$\underline{\zeta}_i \leq \zeta_i \leq \bar{\zeta}_i \quad (4)$$

A modified uniform Dirichlet distribution (Diun) is used in this work²³ to guarantee that the sample corrections sum up to the right-hand side of eq 3 and that individual corrections are within the set bounds. We start with a set of M random variables $\{\zeta_i^*\}$ with joint distribution defined by Diun.²³ Here, the Diun samples obey the following relation.

$$\sum_{i=1}^M \zeta_i^* = 1 \quad (5)$$

To generate the samples for gas molecule corrections using Diun samples, first we have to define a transformation that will preserve the sum of corrections to be equal with the overall thermodynamics. The proposed relationship between Diun random variables and the gas molecule free energy correction is given by the following linear transformation.

$$\zeta_i = a + (b - a)\zeta_i^* \quad (6)$$

Here, $a = \min(\{\zeta_i\}_{i=1 \dots M})$ and $b = \delta - a(M - 1)$. Note that the individual upper bounds $\bar{\zeta}_i$ should be smaller than b . This transformation only ensures that the gas molecule corrections sum up to the right-hand side of eq 3, i.e., that the overall Gibbs' free energy of reaction agrees exactly with NIST data. However, the sampling algorithm needs to take into account that the correction samples generated using this transformation might be larger than the imposed upper bound. As a result, all sample values higher than the upper bound are rejected, and sampling is continued until the desired number of feasible samples is obtained. Gas molecule energy corrections affect all state and transition state energies that contain gas molecules. Given a DFT energy sample \mathbf{y} from the latent variable model previously introduced, a set of corrections $\{\zeta_i\}$ is drawn such that the corrected energy sample $\mathbf{y}_{\text{corrected}}$ matches the overall thermodynamics.

2.3. Propagation of Uncertainty to Quantities of Interest. Consider the turnover frequency (TOF) given by a mapping involving a microkinetic simulation that depends on the corrected DFT energies:

$$\text{TOF} = f(\mathbf{y}_{\text{corrected}}) \quad (7)$$

The following relation gives the probability distribution of TOF induced by the DFT errors and magnitudes of gas molecule corrections.

$$p(\text{TOF}) = \int p(\text{TOF}|\mathbf{y}_{\text{corrected}})p(\mathbf{y}_{\text{corrected}})d\mathbf{y}_{\text{corrected}} \quad (8)$$

Here, a Monte Carlo simulation is proposed to obtain samples from the probability distribution of TOF. A TOF sample is generated as follows. First an energy sample is generated for every intermediate state and transition state according to the

latent variable model. Second, a set of gas molecule corrections is generated using the modified Diun. The corrected energy sample is then used to calculate the rate constants and solve the microkinetic model. The TOF result of the microkinetic model is stored, and a new set of energies is selected; this process is repeated a significant number of times.

Similarly, one can quantify the uncertainty of other QoIs for comparison with experiments such as activation barrier (eV) and reaction orders. The apparent activation barrier is calculated as

$$E_{\text{app}} = k_B T^2 \left(\frac{d \ln(\text{TOF})}{dT} \right)_{P,T} \quad (9)$$

where P is pressure, T is temperature, and k_B is Boltzmann's constant. A range of temperature points and a linear fit are used to calculate E_{app} (eV). Reaction orders, α_i , for each species are calculated as

$$\alpha_i = \left(\frac{d \ln(\text{TOF})}{d \ln(P_i)} \right)_{T, P_{j \neq i}} \quad (10)$$

where P_i (atm) represents the partial pressure of gas species, i .

2.4. Determining the Dominant Catalytic Cycle. The microkinetic model generates a variety of results comparable with experiments. The overall TOF (s^{-1}) as well as individual catalytic cycle TOFs are solved for. Calculating individual catalytic cycle TOF uncertainties allows evaluating the dominant cycle in a reliable manner. The dominant catalytic cycle may be identified using Kullback–Leibler (KL) divergence,^{54,55} which is used in information theory to measure the difference between two probability distributions. KL divergence is defined by

$$D_{\text{KL}}(P \parallel Q) = \int_x p(x) \ln \left(\frac{p(x)}{q(x)} \right) \quad (11)$$

where x is an uncertain variable. p and q are two probabilities for a given value of x . The KL divergence is taken for each catalytic cycle TOF q from the overall TOF p in the case study. A large KL divergence means PDF q is far from PDF p . The KL divergence is non-negative, and it is zero when the two PDF's p and q are identical. A trapezoidal integration in MATLAB, *trapz*, is utilized to evaluate the integral in which the individual catalytic cycle TOF is q and the overall TOF is p . Specifically, a kernel density estimator, MATLAB's *ksdensity*, is used to obtain a set of probabilities and their corresponding TOF values for the trapezoidal integration. The dominant cycle is given by the catalytic cycle TOF that has the smallest KL divergence with respect to the overall TOF. Figure 1 illustrates the workflow of the general framework provided in this section.

3. WGS APPLICATION

Computational and model details of our case study model of the WGS over Pt_8/TiO_2 are available in Ammal et al.²⁴ DFT calculations used the PBE¹⁷ functional to approximate exchange and correlation effects. For a factor analysis,²² three more functionals are used along with PBE to obtain the correlations for a covariance matrix. The additional functionals are the revised Perdew–Burke–Ernzerhof (RPBE),^{18,19} Heyd–Scuseria–Ernzerhof (HSE),²⁰ and M06L²¹ which is a Minnesota functional. PBE uses a GGA which utilizes fundamental parameters apart from local spin density parameters to approximate the exchange–correlation energy in Kohn–Sham DFT and is

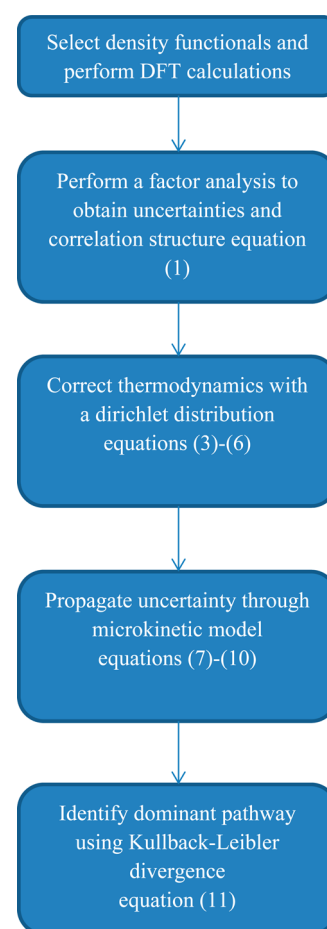


Figure 1. Flowchart of general framework for uncertainty quantification of computational catalysis results presented in section 2.

considered a universally good functional.⁵⁶ RPBE belongs to the class of GGA functionals that has been optimized for chemisorption energetics on transition metal surfaces to overcome the observed limitation of PBE to often overestimate the adsorption energy of various small molecules.¹⁸ Interestingly for this study, it has been proposed that PBE overestimates the binding of CO and H, while RPBE misrepresents the energy of species with OCO backbone.⁹ All of these species are of significant importance for the WGS. Next, HSE is a computationally efficient hybrid exchange density functional based on the PBE0 (or sometimes called PBEh) functional that contains 25% exact exchange.^{20,57} For nonmetallic systems, transition state energies, and (of particular importance for this study) oxygen vacancy formation energies, it has been argued that the inclusion of exact exchange is important to describe the electronic structure correctly.^{16,58} Finally, we also performed calculations with the M06L meta-GGA DFT functional that is computationally faster than hybrid exchange functionals (though significantly slower than pure GGA functionals) and that has been optimized to predict good energetics (including energy barriers and medium range dispersion interactions) for transition metal, inorganic, and organometallic systems.

Climbing image nudge elastic band and dimer methods were used to obtain transition states^{24,59,60} at the PBE level of theory as reported previously. For all other DFT functionals we only performed single point energy calculations on the PBE optimized geometries. A full optimization with each DFT functional would be preferred; however, the computational expense would be

Table 1. Four Functional Results for the Gibbs' Free Energy of Various States of Our Model Originally Reported by Ammal et al.^{24,a}

intermediate or transition state	G (eV)			
	PBE	RPBE	HSE	M06L
* _{Pt} -O _{int} + 2CO(g) + H ₂ O(g)	0.000	0.000	0.000	0.000
* _{Pt} -O _{int} + 2CO(g) + H ₂ O(g) → CO _{Pt} -O _{int} + CO(g) + H ₂ O(g)	0.000	0.000	0.000	0.000
CO _{Pt} -O _{int} + CO(g) + H ₂ O(g)	0.182	0.273	0.365	0.768
CO _{Pt} -O _{int} + CO(g) + H ₂ O(g) → CO ₂ (Pt-int) + CO(g) + H ₂ O(g)	0.733	0.951	0.940	1.399
CO ₂ (Pt-int) + CO(g) + H ₂ O(g)	0.605	0.892	0.748	0.957
CO ₂ (Pt-int) + CO(g) + H ₂ O(g) → * _{Pt} -V _{int} + CO ₂ (g) + CO(g) + H ₂ O(g)	1.829	1.976	2.803	2.095
* _{Pt} -V _{int} + CO(g) + H ₂ O(g) + CO ₂ (g)	0.489	0.408	0.868	1.074
* _{Pt} -V _{int} + CO(g) + H ₂ O(g) + CO ₂ (g) → * _{Pt} -2OH _{int} + CO(g) + CO ₂ (g)	1.020	1.263	2.044	1.395
* _{Pt} -2OH _{int} + CO(g) + CO ₂ (g)	0.121	0.462	0.667	-0.353
* _{Pt} -2OH _{int} + CO(g) + CO ₂ (g) → H _{Pt} -OH _{int} + O _{int} + CO(g) + CO ₂ (g)	0.697	0.905	1.147	1.349
H _{Pt} -OH _{int} + * _{Pt} + CO(g) + CO ₂ (g)	0.374	0.507	0.736	0.842
H _{Pt} -OH _{int} + * _{Pt} + CO(g) + CO ₂ (g) → 2H _{Pt} -O _{int} + CO(g) + CO ₂ (g)	1.036	1.424	1.853	1.339
2H _{Pt} -O _{int} + CO(g) + CO ₂ (g)	-0.097	0.245	0.309	0.175
2H _{Pt} -O _{int} + CO(g) + CO ₂ (g) → * _{Pt} -O _{int} + * _{Pt} + H ₂ (g) + CO(g) + CO ₂ (g)	-0.097	0.245	0.309	0.175
* _{Pt} -O _{int} + * _{Pt} + H ₂ (g) + CO(g) + CO ₂ (g)	-0.667	-0.622	-0.468	-0.803
CO ₂ (Pt-int) + CO(g) + H ₂ O(g) → CO _{Pt} -CO ₂ (int) + H ₂ O(g)	0.605	0.892	0.748	0.957
CO _{Pt} -CO ₂ (int) + H ₂ O(g)	0.540	0.780	0.879	1.318
CO _{Pt} -CO ₂ (int) + H ₂ O(g) → CO _{Pt} -V _{int} + CO ₂ (g) + H ₂ O(g)	1.272	1.323	1.851	2.306
CO _{Pt} -V _{int} + O _{int} + H ₂ O(g) + CO ₂ (g)	-0.305	-0.263	0.607	0.497
CO _{Pt} -V _{int} + O _{int} + H ₂ O(g) + CO ₂ (g) → CO _{Pt} -2OH _{int} + CO ₂ (g)	0.924	1.050	1.496	1.958
CO _{Pt} -2OH _{int} + O _s + CO ₂ (g)	0.313	0.673	0.957	0.570
CO _{Pt} -2OH _{int} + O _s + CO ₂ (g) → CO _{Pt} -OH _{int} -O _{int} -OH _s + CO ₂ (g)	0.771	1.111	1.206	1.651
CO _{Pt} -OH _{int} -O _{int} -OH _s + O _{int} + CO ₂ (g)	0.379	0.571	0.658	1.094
CO _{Pt} -OH _{int} -O _{int} -OH _s + O _{int} + CO ₂ (g) → CO _{Pt} -OH _{int} -O _{int} -OH _{int} + O _s + CO ₂ (g)	0.869	0.949	1.258	1.920
CO _{Pt} -OH _{int} -O _{int} -OH _{int} + * _{Pt} + CO ₂ (g)	0.236	0.517	0.777	0.514
CO _{Pt} -OH _{int} -O _{int} -OH _{int} + * _{Pt} + CO ₂ (g) → CO _{Pt} -OH _{int} -O _{int} -H _{Pt} + O _{int} + CO ₂ (g)	0.262	0.397	0.648	1.344
CO _{Pt} -OH _{int} -O _{int} -H _{Pt} + * _{Pt} + CO ₂ (g)	-0.364	-0.097	-0.174	0.219
CO _{Pt} -OH _{int} -O _{int} -H _{Pt} + * _{Pt} + CO ₂ (g) → CO _{Pt} -O _{int} -2H _{Pt} + O _{int} + CO ₂ (g)	-0.034	0.148	0.296	1.106
CO _{Pt} -O _{int} -2H _{Pt} + CO ₂ (g)	-0.761	-0.483	-0.612	0.102
CO _{Pt} -O _{int} -2H _{Pt} + CO ₂ (g) → CO _{Pt} -O _{int} + 2* _{Pt} + CO ₂ (g) + H ₂ (g)	-0.761	-0.483	-0.612	0.102
CO _{Pt} -O _{int} + 2* _{Pt} + CO ₂ (g) + H ₂ (g)	-0.485	-0.349	-0.103	-0.034

^aA temperature of 523 K has been used for the free energy calculations at a reference pressure of 1 atm for each gas molecule. All energies are with reference to state: *_{Pt}-O_{int} + 2CO(g) + H₂O(g). Transition states are symbolized by the reaction equation they belong to.

significantly larger (particularly for the HSE and M06L calculations) such that we refrained from using such an approach here. Transition state theory within the harmonic approximation and for adsorption processes collision theory were used to calculate reaction rate constants. For the vibrational partition function and zero-point energy (ZPE) correction we used the numbers from the PBE functional where we shifted all vibrational frequencies smaller than 100 cm⁻¹ to 100 cm⁻¹ to avoid having large entropy contributions from very small frequencies. The vibrational frequencies are computed by numerical differentiation. A 0.01 Å displacement was used for all vibrational frequency calculations. There are four catalytic cycles considered in this case study. The first catalytic cycle is the classical redox catalytic cycle.^{24,61,62} The second catalytic cycle is the CO-promoted redox catalytic cycle.²⁴ Also, the final two pathways are the formate and carboxyl pathways with redox regeneration that we predicted previously to be of no importance.²⁴ All elementary reaction steps in these four pathways have been included in a microkinetic model as described previously.²⁴ The *gsl* (GNU Scientific Library) linear algebra library in C++ is used to solve the set of steady state reactor equations.

Next, we placed Gaussian PDFs representing the DFT energy uncertainty and their correlation on all intermediate and

transition states. Transition states from collision theory are given a PDF with correction with a mean of +0.075 eV and a standard deviation of 0.075 eV. This mean correction suggests collision theory tends to underestimate transition state energies for adsorption. The selection of a standard deviation covers a sticking coefficient from zero to one where at 500 K a correction of 0.075 eV corresponds to a sticking coefficient of 0.175.

The standard deviations of all other intermediate and transition states are obtained by the results of our factor analysis²² of the four functionals. Only one factor is used to explain the correlation structure between DFT energies. The DFT calculations with four functional results are listed in Table 1. The standard deviations are listed in Table 2. All standard deviations are from the element-wise square root of the diagonal of the covariance matrix obtained through factor analysis. Free energies of reaction and activation barriers which vary more widely are reflected by relatively larger standard deviations in Table 2.

The DFT-derived free energy pathways for this system including all intermediate and transition states need to end 0.40 eV higher to match the overall reaction thermodynamics (NIST value, $\Delta G_{\text{rxn}}^{\text{WGS}}(523 \text{ K}) = -0.24 \text{ eV}$; average DFT, $\Delta G_{\text{rxn}}^{\text{WGS}}(523 \text{ K}) = -0.64 \text{ eV}$; PBE, $\Delta G_{\text{rxn}}^{\text{WGS}}(523 \text{ K}) = -0.67 \text{ eV}$; RPBE,

Table 2. Standard Deviations of Gaussian Uncertainties of Gibbs' Free Energies Obtained by Factor Analysis and Setting Adsorption/Desorption Uncertainties^a

intermediate or transition state	standard deviation (eV)
* _{Pt} -O _{int} + 2CO(g) + H ₂ O(g)	0.000
* _{Pt} -O _{int} + 2CO(g) + H ₂ O(g) → CO _{Pt} -O _{int} + CO(g) + H ₂ O(g)	0.075
CO _{Pt} -O _{int} + CO(g) + H ₂ O(g)	0.224
CO _{Pt} -O _{int} + CO(g) + H ₂ O(g) → CO ₂ (Pt-int) + CO(g) + H ₂ O(g)	0.243
CO ₂ (Pt-int) + CO(g) + H ₂ O(g)	0.136
CO ₂ (Pt-int) + CO(g) + H ₂ O(g) → * _{Pt} -V _{int} + CO ₂ (g) + CO(g) + H ₂ O(g)	0.374
* _{Pt} -V _{int} + CO(g) + H ₂ O(g) + CO ₂ (g)	0.273
* _{Pt} -V _{int} + CO(g) + H ₂ O(g) + CO ₂ (g) → * _{Pt} -2OH _{int} + CO(g) + CO ₂ (g)	0.379
* _{Pt} -2OH _{int} + CO(g) + CO ₂ (g)	0.386
* _{Pt} -2OH _{int} + CO(g) + CO ₂ (g) → H _{Pt} -OH _{int} + O _{int} + CO(g) + CO ₂ (g)	0.246
H _{Pt} -OH _{int} + * _{Pt} + CO(g) + CO ₂ (g)	0.184
H _{Pt} -OH _{int} + * _{Pt} + CO(g) + CO ₂ (g) → 2H _{Pt} -O _{int} + CO(g) + CO ₂ (g)	0.292
2H _{Pt} -O _{int} + CO(g) + CO ₂ (g)	0.155
2H _{Pt} -O _{int} + CO(g) + CO ₂ (g) → * _{Pt} -O _{int} + * _{Pt} + H ₂ (g) + CO(g) + CO ₂ (g)	0.075
* _{Pt} -O _{int} + * _{Pt} + H ₂ (g) + CO(g) + CO ₂ (g)	0.119
CO ₂ (Pt-int) + CO(g) + H ₂ O(g) → CO _{Pt} -CO ₂ (int) + H ₂ O(g)	0.136
CO _{Pt} -CO ₂ (int) + H ₂ O(g)	0.282
CO _{Pt} -CO ₂ (int) + H ₂ O(g) → CO _{Pt} -V _{int} + CO ₂ (g) + H ₂ O(g)	0.423
CO _{Pt} -V _{int} + O _{int} + H ₂ O(g) + CO ₂ (g)	0.420
CO _{Pt} -V _{int} + O _{int} + H ₂ O(g) + CO ₂ (g) → CO _{Pt} -2OH _{int} + CO ₂ (g)	0.407
CO _{Pt} -2OH _{int} + O _s + CO ₂ (g)	0.231
CO _{Pt} -2OH _{int} + O _s + CO ₂ (g) → CO _{Pt} -OH _{int} -O _{int} -OH _s + CO ₂ (g)	0.314
CO _{Pt} -OH _{int} -O _{int} -OH _s + O _{int} + CO ₂ (g)	0.262
CO _{Pt} -OH _{int} -O _{int} -OH _s + O _{int} + CO ₂ (g) → CO _{Pt} -OH _{int} -O _{int} -OH _{int} + O _s + CO ₂ (g)	0.414
CO _{Pt} -OH _{int} -O _{int} -OH _{int} + * _{Pt} + CO ₂ (g)	0.191
CO _{Pt} -OH _{int} -O _{int} -OH _{int} + * _{Pt} + CO ₂ (g) → CO _{Pt} -OH _{int} -O _{int} -H _{Pt} + O _{int} + CO ₂ (g)	0.417
CO _{Pt} -OH _{int} -O _{int} -H _{Pt} + * _{Pt} + CO ₂ (g)	0.210
CO _{Pt} -OH _{int} -O _{int} -H _{Pt} + * _{Pt} + CO ₂ (g) → CO _{Pt} -O _{int} -2H _{Pt} + O _{int} + CO ₂ (g)	0.436
CO _{Pt} -O _{int} -2H _{Pt} + CO ₂ (g)	0.327
CO _{Pt} -O _{int} -2H _{Pt} + CO ₂ (g) → CO _{Pt} -O _{int} + 2* _{Pt} + CO ₂ (g) + H ₂ (g)	0.075
CO _{Pt} -O _{int} + 2* _{Pt} + CO ₂ (g) + H ₂ (g)	0.398

^aStandard deviations correspond to how varied the four functional results (free energies) listed in Table 1 are. All energies are with reference to state: *_{Pt}-O_{int} + 2CO(g) + H₂O(g). Transition states are symbolized by the reaction equation they belong to.

$\Delta G_{\text{rxn}}^{\text{WGS}}(523 \text{ K}) = -0.62 \text{ eV}$; HSE, $\Delta G_{\text{rxn}}^{\text{WGS}}(523 \text{ K}) = -0.47 \text{ eV}$; M06L, $\Delta G_{\text{rxn}}^{\text{WGS}}(523 \text{ K}) = -0.80 \text{ eV}$. A four-dimensional modified Diun is used to obtain four corrections ζ_{CO_2} , ζ_{H_2} , $\zeta_{\text{H}_2\text{O}}$, ζ_{CO} to gas molecule energies for CO₂, H₂, H₂O and CO, respectively. The following constraints are imposed on the four corrections.

$$\begin{aligned} \zeta_{\text{CO}_2} + \zeta_{\text{H}_2} + \zeta_{\text{H}_2\text{O}} + \zeta_{\text{CO}} &= \delta \\ -0.4 &\leq \zeta_{\text{CO}_2}, \zeta_{\text{CO}} \leq 0.4 \\ -0.2 &\leq \zeta_{\text{H}_2}, \zeta_{\text{H}_2\text{O}} \leq 0.2 \end{aligned} \quad (12)$$

We choose relatively large bounds on uncertainty of $\pm 0.4 \text{ eV}$ for CO and CO₂ due to the known challenges of DFT to

predict both CO and CO₂ equally well,⁴ while we choose relatively small uncertainty bounds for H₂ and H₂O of $\pm 0.2 \text{ eV}$. It can be argued that our uncertainty of ± 0.4 and $\pm 0.2 \text{ eV}$, respectively, is too large; however, we advise readers to not underestimate uncertainties and note that the overall error of the average of our four functionals is 0.40 eV for the WGS at $T = 523 \text{ K}$. The following linear transformation is used to alter the Diun distribution:

$$\begin{aligned} \zeta_{\text{CO}_2} &= a + (b - a)\zeta_{\text{CO}_2}^* \\ \zeta_{\text{H}_2} &= a + (b - a)\zeta_{\text{H}_2}^* \\ \zeta_{\text{H}_2\text{O}} &= a + (b - a)\zeta_{\text{H}_2\text{O}}^* \\ \zeta_{\text{CO}} &= a + (b - a)\zeta_{\text{CO}}^* \\ a &= -0.4 \\ b &= 0.40 - 3a \end{aligned} \quad (13)$$

This defines a four-dimensional Dirichlet distribution, called a simplex, where each correction is bounded below by a and above by b . This simplex is cut in order to provide the ranges of eq 12. Figure 2 illustrates samples from eq 13 following the

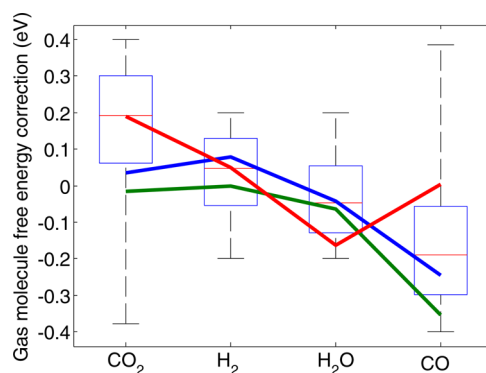


Figure 2. Box plot of gas molecule free energy corrections with three samples shown on top. The overall reaction thermodynamics may be expressed as $\sum_i \nu_i G_i$ where ν_i is the stoichiometric coefficient of gas molecule i . Overall reaction Gibbs' free energies are $\Delta G^{\text{error}} = G_{\text{CO}_2}^{\text{error}} + G_{\text{H}_2}^{\text{error}} - G_{\text{H}_2\text{O}}^{\text{error}} - G_{\text{CO}}^{\text{error}}$. The three samples illustrate that ΔG^{error} is $+0.40 \text{ eV}$. For example, a possible sample from the four-dimensional Dirichlet PDF is $0.15, 0.07, -0.01, -0.17 \text{ (eV)}$ for CO₂, H₂, H₂O, and CO because it falls within the constraints and sums (using stoichiometric coefficients) to $+0.40 \text{ eV}$. If a three-dimensional slice were taken of this four-dimensional space the samples would appear uniformly spread.

ranges of eq 12 as a box plot. Three samples are plotted on top of the box plot to illustrate that each sample of gas DFT energy corrections sums to 0.40 eV .

Another coding strategy is included to ensure that energies in the UQ simulation are physical. Transition states used in transition state theory are set so that no activation barriers may be negative. This is accomplished by constraining transition state free energies to be at least equal to product or reactant state free energies, whichever is greater.

4. RESULTS AND DISCUSSION

Figure 3a,b shows free energy diagrams which visualize the four functional results of the classical redox pathway and the CO-promoted redox pathway on our edge interface site model

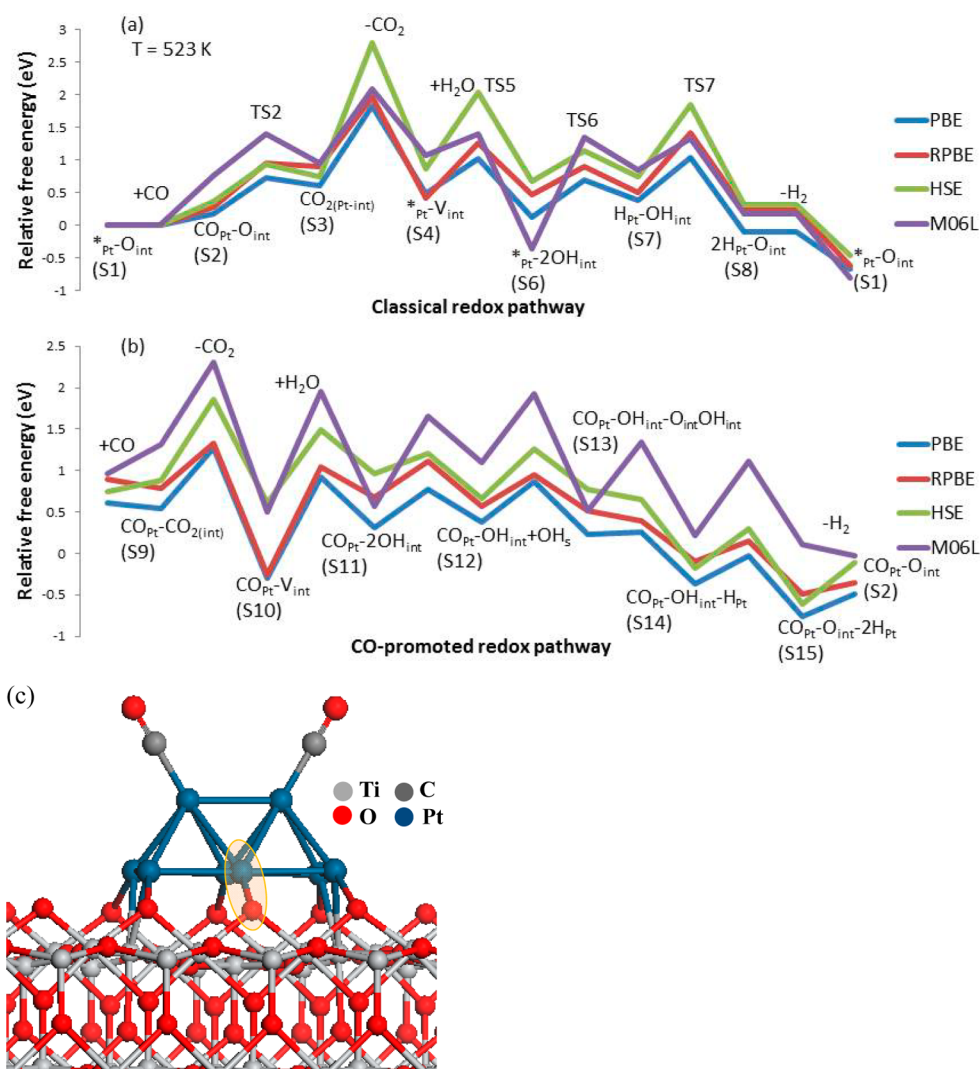


Figure 3. (a–c) Free energy diagrams of the classical redox pathway (a) and CO-promoted redox pathway (b) computed by four DFT functionals at a reaction temperature of 523 K. Part c is a picture of the active edge interface site model.

pictured in Figure 3c.²⁴ We observe that PBE and RPBE free energies are relative to the HSE and M06L results and are quite similar. Also, HSE and M06L free energies can deviate from the PBE and RPBE results by as much as 1 eV, particularly for transition states involving the formation of an oxygen vacancy in the TiO_2 support. Figure 4a shows the 95% confidence interval based upon average DFT energies. The 95% confidence interval is the mean of the four functionals plus or minus 2 standard deviations, indicated by dashed lines. The mean of the four functionals is the expected value, the solid line. Figure 4b shows uncertainties from the factor analysis with thermodynamics (Gibbs' free energy) correction for the gas molecules. The final and initial Gibbs' free energies are entirely certain in Figure 4b (the initial state is the reference energy, and the final state is given by the reaction thermodynamics). In other words, the mean of the four functionals undergoes gas-phase corrections which correct the Gibbs' free energy of reaction to the NIST value which is assumed to have no uncertainty (the NIST uncertainty is significantly smaller than the DFT uncertainty).

Figure 5 displays overall TOF results and TOF results of individual catalytic cycles including constraints on (gas molecule) thermodynamics and correlations. The vertical

dashed line is the experimental observation from Kalamaras et al.⁴⁴ The dominant catalytic cycle is predicted to be the CO-promoted redox pathway at these reaction conditions ($P_{\text{CO}} = 0.03$ atm, $P_{\text{H}_2\text{O}} = 0.1$ atm, $P_{\text{CO}_2} = 0.06$ atm, $P_{\text{H}_2} = 0.2$ atm, $T = 523$ K). This observation is in agreement with our previous conclusions about the dominant reaction mechanism on Pt_8/TiO_2 interface sites.²⁴ Figure 5 displays evidence of this trend by visual inspection that the probability density of the TOF of the CO-promoted redox pathway essentially lines up with the overall TOF probability density. This lining up signifies that there is hardly any difference between overall TOF and CO-promoted TOF. Although the classical redox pathway is not lined up with the overall TOF, the classical redox pathway is able to reach the experimentally observed TOF. It is evident that even when considering uncertainty, the formate and carboxyl pathway cannot lead to a significant WGS rate at edge sites of a Pt cluster on TiO_2 (110), and we can exclude the possibility that the WGS proceeds through these pathways at interface edge sites.

To statistically formalize the importance of one pathway over another, the KL divergence is taken of the CO-promoted redox catalytic cycle TOF and the classical redox catalytic cycle TOF using the overall TOF as a reference over the temperature

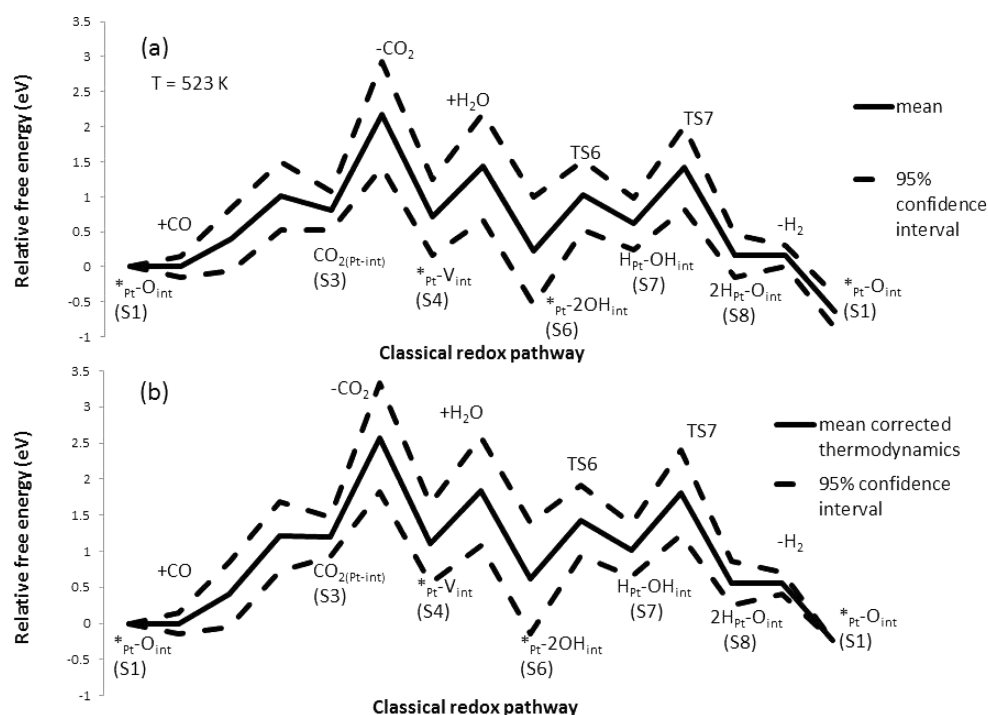


Figure 4. (a, b) Relative Gibbs' free energy diagram of the classical redox cycle with 95% confidence intervals from factor analysis. Part b shows the dirichlet PDF results correcting the thermodynamics (Gibbs' free energy of reaction) to the NIST value at the end of the catalytic cycle.

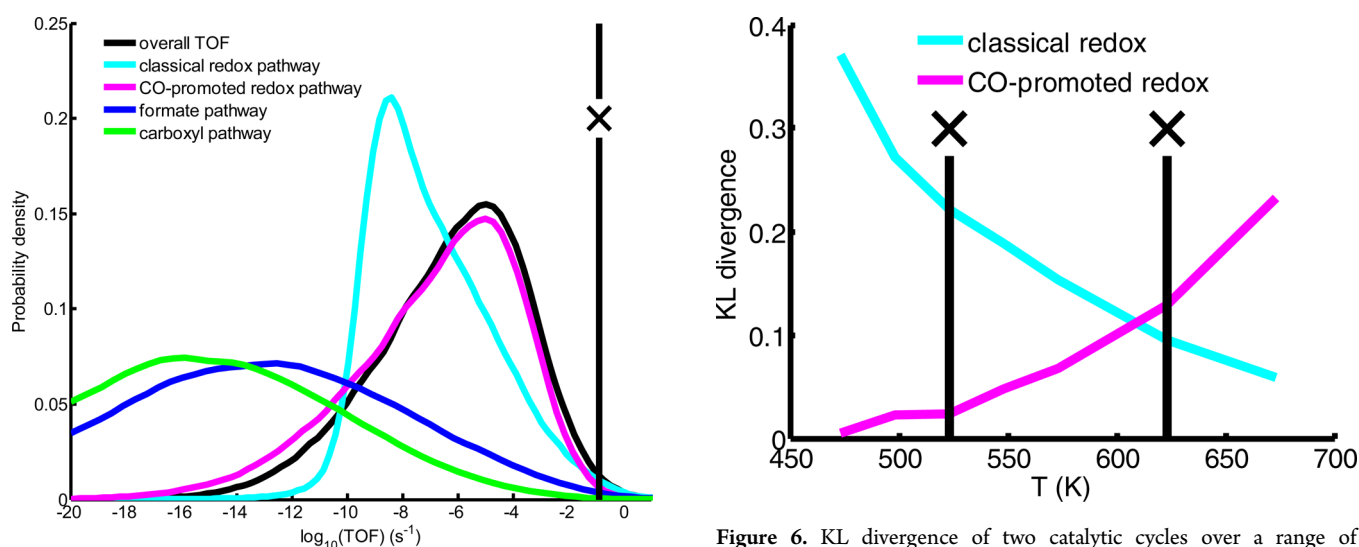


Figure 5. Probability density of TOFs calculated for different pathways of the WGS at Pt/TiO₂ interface sites. At least 70 000 samples are used. 470 000 samples are used to test for convergence of the 70 000 sample results. No change is observed, and 70 000 sample results are plotted: $P_{\text{CO}} = 0.03$ atm, $P_{\text{H}_2\text{O}} = 0.1$ atm, $P_{\text{CO}_2} = 0.06$ atm, $P_{\text{H}_2} = 0.2$ atm, $T = 523$ K. The straight line with X corresponds to the experimental results from Kalamaras et al.⁴⁴

range 473–673 K as shown in Figure 6. The raw numerical data from the model must be transformed and entered into the equation for KL divergence, eq 11. A smaller KL divergence reveals that the individual catalytic cycle TOF is closer to the overall TOF. As a note, any plotting software/language which has kernel density estimation capability could create the plots which MATLAB creates. As a clarification, MATLAB is only used for plotting the data in this work. QUESO, gsl, and other C++ libraries are used to build and run the model for the

Figure 6. KL divergence of two catalytic cycles over a range of temperatures. A smaller KL divergence suggests dominance. $D_{\text{KL}}(P||Q) = \int_x p(x) \log_2 \left(\frac{p(x)}{q(x)} \right)$. Two temperatures at 523 and 623 K indicated by straight line with X are selected for discussion.

generation of data. Python is used to obtain confidence intervals and correlation structure by factor analysis.⁶³ At a reaction temperature of 523 K, the KL divergence of the CO-promoted redox catalytic cycle is 0.0237 which is significantly less than the 0.2211 KL divergence of the classical redox cycle. Therefore, we have high confidence that the CO-promoted catalytic cycle is indeed dominant. The CO-promoted redox catalytic cycle dominates within the temperature range of 473 K to at least 600 K, at which temperature we observe slowly a change in reaction mechanism (the KL for both pathways becomes equivalent at 608 K). This change in dominant catalytic cycle is likely due to the effect of temperature on the CO binding energy and therefore the surface coverage fraction of the extra

adsorbed CO molecule needed in the CO-promoted cycle. We also analyzed the reaction pathways at a temperature of 623 K, and Figure 7 illustrates the PDFs of the two cycles and overall

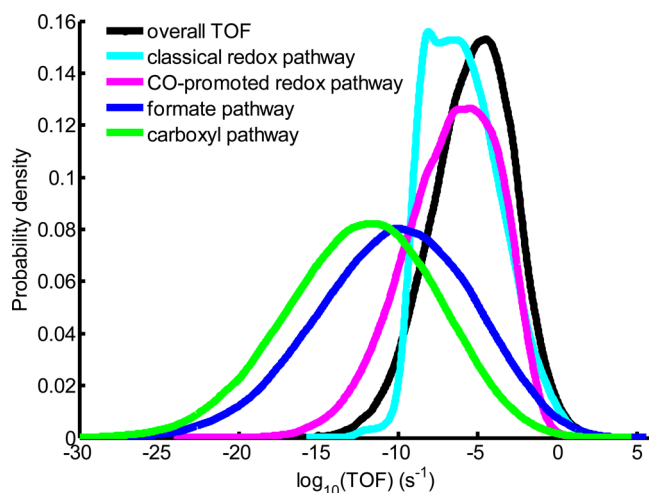


Figure 7. Probability density of overall TOF and individual catalytic cycle TOF of various pathways at $P_{\text{CO}} = 0.03$ atm, $P_{\text{H}_2\text{O}} = 0.1$ atm, $P_{\text{CO}_2} = 0.06$ atm, $P_{\text{H}_2} = 0.2$ atm, and $T = 623$ K. The catalytic cycles have moved closer together with an increase in temperature of 100 K relative to Figure 4 as indicated by the KL divergence in Figure 5.

TOF at $T = 623$ K. Here, the classical redox TOF is very close to overall TOF, and it becomes very challenging to identify the dominant reaction mechanism by visual inspection. However, it is still apparent that the formate and carboxyl mechanism lead to small TOFs, and these pathways are not relevant for our Pt/TiO₂ catalyst model.

Next, we analyzed our predicted apparent activation barrier and reaction orders. The apparent activation barrier of our active site model captures the experiment as evidenced by Figure 8 although it has wide uncertainty of over 2.5 eV

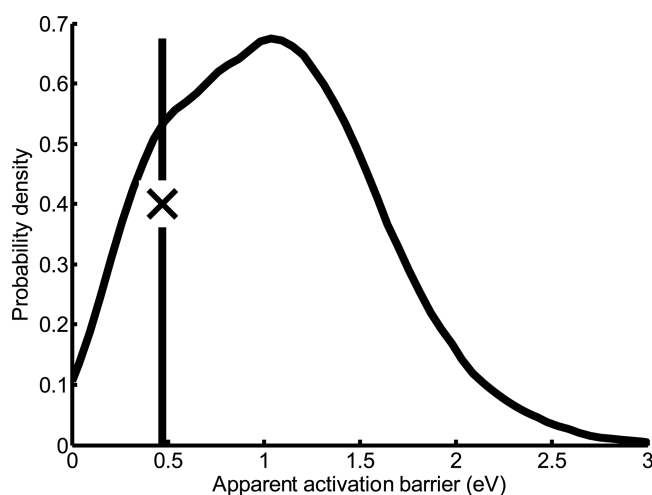


Figure 8. Probability density of the apparent activation barrier obtained at $P_{\text{CO}} = 0.03$ atm, $P_{\text{H}_2\text{O}} = 0.1$ atm, $P_{\text{CO}_2} = 0.06$ atm, $P_{\text{H}_2} = 0.2$ atm, and $T = 473$ – 623 K. Experimental data point from Kalamaras et al.⁴⁴ is symbolized by a straight line with x.

(the apparent activation barrier has been obtained in the temperature interval from 473 to 673 K). This wide uncertainty is a result of the significant uncertainty of the DFT predicted

Gibbs' free energies (Figures 3 and 4), particularly for processes involving the formation of an oxygen vacancy on the TiO₂ surface. In other words, agreement between experimental and computational apparent activation barriers is not a sufficient measure to validate the computational model. We note however that there is a correlation between the TOF and apparent activation barrier such that when considering only the parameter space with results that are close to the experimental TOF, the uncertainty in apparent activation barrier is significantly reduced and relatively close to the experimentally observed barrier. Finally, Figure 9 shows reaction order PDFs

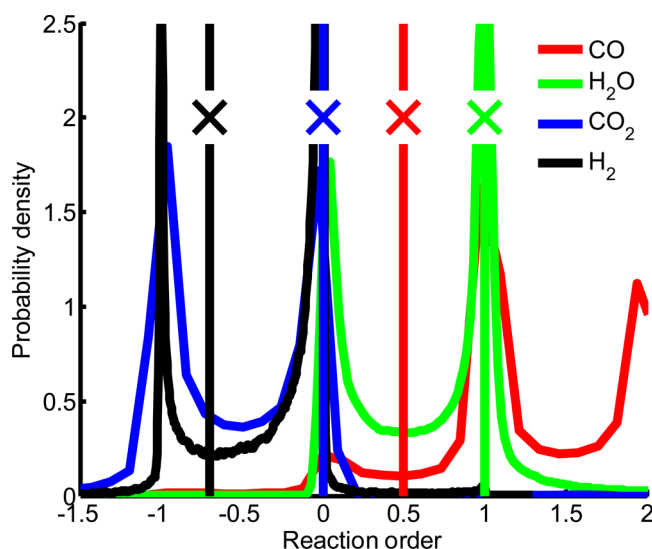


Figure 9. Probability density of various reaction orders. All experimental data from Kalamaras et al.⁴⁴ (symbolized by straight lines with x) are captured by their PDFs. Reaction orders have been determined at the following reaction conditions: $P_{\text{CO}} = 0.03$ (0.02–0.08) atm, $P_{\text{H}_2\text{O}} = 0.1$ (0.05–0.1) atm, $P_{\text{CO}_2} = 0.06$ (0.02–0.1) atm, $P_{\text{H}_2} = 0.2$ (0.05–0.4) atm, and $T = 523$ K.

with experiments reported in the literature at reaction conditions of $P_{\text{CO}} = 0.03$ atm, $P_{\text{H}_2\text{O}} = 0.1$ atm, $P_{\text{CO}_2} = 0.06$ atm, $P_{\text{H}_2} = 0.2$ atm, $T = 523$ K.¹⁵ The experimental H₂ reaction order is reasonably captured well by its PDF. H₂ and in fact all reaction orders display some degree of bimodality due to the large uncertainty in the DFT energies which leads to various states being rate controlling in different samples; for example, the reaction order for H₂O can be zero or one (or anywhere in between) when the uncertainty in DFT is accounted. Also, the CO reaction order PDF is able to capture the experiment, and the H₂O reaction order PDF captures the experimental observation with most probability lying between 0 and 1. Finally, the CO₂ reaction order PDF captures the experimental value at its mode at 0. Overall, we observe wide ranges of predicted reaction orders with bimodality that capture the experimental data. In other words, our model based on DFT and transition state theory captures all experimental data (indeed there exist free energy corrections that lead to kinetic parameters that all agree concurrently with experimental data); however, the PDFs for TOF, apparent activation barrier, and reaction orders are relatively wide such that DFT is still challenged with predicting the absolute activity of active sites at the interface of a transition metal particle and a reducible oxide support. Encouragingly though the prediction of the dominant reaction pathway/mechanism can be made with high confidence

for even complex active site models. Clearly, the formate and carboxyl mechanisms play no role at the investigated active site model, and the CO-promoted redox pathway is the dominant reaction mechanism at low temperatures.

5. CONCLUSIONS

A general framework has been proposed for uncertainty quantification of computational catalysis studies. The usefulness of correlating DFT uncertainties using a factor analysis has been shown. A Dirichlet distribution has been used to perform a thermodynamics correction. Case study results have been presented, and they suggest DFT is accurate for producing relative results such as the dominant catalytic cycle. Kullback–Leibler divergence has been presented for identification of the dominant catalytic cycle. For the case study of the water–gas shift reaction at the edge interface site of our Pt₈/TiO₂ catalyst model, a combination of generalized-gradient approximation functionals, a hybrid functional, and a Minnesota functional have been entered into the factor analysis for DFT uncertainty and correlation structure among states and transition states. The experimental data of TOF, apparent activation barrier, and reaction orders of CO, H₂O, CO₂, and H₂ all fall within the uncertainty due to DFT illustrating that our active site model might capture all relevant characteristics of the active sites in the experimental catalyst. Two catalytic cycles, formate and carboxylate, have been ruled out from playing a dominant role in TOF. Furthermore, the Kullback–Leibler divergence formalizes that the CO-promoted cycle is dominant over the classical redox cycle. However, above 600 K the two cycles become equivalent in dominance as measured by Kullback–Leibler divergence. A four-dimensional Dirichlet distribution has been applied to the gas molecules CO, H₂O, CO₂, and H₂, which corrects the overall Gibbs' free energy of reaction to NIST data; i.e., the reaction is less thermodynamically downhill with the Dirichlet distribution correction. Uncertainty quantification *forward problems* such as the general framework and case study communicated here are recommended for other computational catalysis models in order to account for the inexact nature of DFT and can of course be extended to include other uncertainties in the model if they are believed to be of high importance such as lateral interactions and limitations of harmonic transition state theory. In future work, a significant reduction in uncertainty may be accomplished by using Bayesian statistical tools to perform an *inverse problem* to learn about DFT energies and their correlation from experiments.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: TEREJANU@cse.sc.edu. Phone: 803-777-5872.

*E-mail: heyden@cec.sc.edu. Phone: 803-777-5025.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to acknowledge the National Science Foundation (NSF) CAREER award Grant NSF CBET-1254352 and the 2014 Eastman fellowship. The US Department of Energy facilities located at the National Energy Research Scientific Computing Center (NERSC) and at EMSL, located at Pacific Northwest National Laboratory, are gratefully acknowledged for their computing resources used in this work. Furthermore, a portion of this research was performed with

XSEDE resources provided by the National Institute for Computational Sciences (NICS), San Diego Supercomputer Center (SDSC), and Texas Advanced Computing Center (TACC) under Grant TG-CTS090100.

LIST OF SYMBOLS

N	total number of density functionals; $N = 4$ in this work
\mathbf{y}	vector of observed DFT energies for intermediate and transition states; in this work, the length of \mathbf{y} is 31, and \mathbf{Y} is the corresponding matrix with N functionals; with thermodynamics correction a subscript is added, i.e., $\mathbf{y}_{\text{corrected}}$
\mathbf{z}	vector of independent factors/latent variables that are normally distributed
\mathbf{W}	matrix containing factor loadings that capture the correlations between observables
\mathbf{e}	vector of zero-mean Gaussian distributed noise $\sim N(0, \Psi)$
Ψ	diagonal covariance matrix that contains <i>specific variances</i>
μ	vector of means of DFT Gibbs' free energies
Σ	covariance matrix containing correlation structure
M	total number of gas molecules
ζ_i	energy correction for gas molecule i
δ	thermodynamics correction to be made before the inclusion of uncertainty
$\underline{\zeta}_i$	lower bound for ζ_i
$\overline{\zeta}_i$	upper bound for ζ_i
ζ_i^*	ζ_i prior to transformation
a	used for transformation of ζ_i^* to ζ_i
b	used for transformation of ζ_i^* to ζ_i
$\mathbf{y}_{\text{corrected}}$	corrected energy sample
E_{app}	apparent activation energy (eV)
α_i	reaction orders for each gas species
k_B	Boltzmann constant
T	temperature
P	pressure
$D_{\text{KL}}(P Q)$	Kullback–Leibler divergence ⁵⁵ of PDF P from PDF Q
x	an uncertain variable; in this study x is $\log_{10}(\text{TOF})$
p, q	two separate probabilities for a given value of x

REFERENCES

- (1) Hansgen, D. A.; Vlachos, D. G.; Chen, J. G. Using First Principles to Predict Bimetallic Catalysts for the Ammonia Decomposition Reaction. *Nat. Chem.* **2010**, *2*, 484–489.
- (2) Studt, F.; Pedersen, F. A.; Bligaard, T.; Sørensen, R. Z.; Christensen, C. H.; Nørskov, J. K. Identification of Non-Precious Metal Alloy Catalysts for Selective Hydrogenation of Acetylene. *Science* **2008**, *320*, 1320–1322.
- (3) Honkala, K.; Hellman, A.; Remediakis, I. N.; Logadottir, A.; Carlsson, A.; Dahl, S.; Christensen, C. H.; Nørskov, J. K. Ammonia Synthesis from First-Principles Calculations. *Science* **2005**, *307*, 555–558.
- (4) Mortensen, J. J.; Kaasbjerg, K.; Frederiksen, S. L.; Nørskov, J. K.; Sethna, J. P.; Jacobsen, K. W. Bayesian Error Estimation in Density-Functional Theory. *Phys. Rev. Lett.* **2005**, *95*, 216401.
- (5) Cramer, C. J. *Essentials of Computational Chemistry*, 2nd ed.; John Wiley and Sons, Inc.: Hoboken, NJ, 2007.
- (6) Hanke, F. Sensitivity Analysis and Uncertainty Calculation for Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1424–1430.

- (7) Sutton, J. E.; Vlachos, D. G. Error Estimates in Semi-Empirical Estimation Methods of Surface Reactions. *J. Catal.* **2013**, *297*, 202–216.
- (8) Sutton, J. E.; Panagiotopoulou, P.; Verykios, X. E.; Vlachos, D. G. Combined DFT, Microkinetic, and Experimental Study of Ethanol Steam Reforming on Pt. *J. Phys. Chem. C* **2013**, *117*, 4691–4706.
- (9) Peterson, A. A.; Pedersen, F. A.; Studt, F.; Rossmeisl, J.; Nørskov, J. K. How Copper Catalyzes the Electroreduction of Carbon Dioxide into Hydrocarbon Fuels. *Energy Environ. Sci.* **2010**, *3*, 1311–1315 Supporting Information.
- (10) Afeefy, H.; Liebman, J.; Stein, S. *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; National Institute of Standards and Technology: Gaithersburg, MD, 2010.
- (11) Petzold, V.; Bligaard, T. K.; Jacobsen, W. Construction of New Electronic Density Functionals with Error Estimation Through Fitting. *Top. Catal.* **2012**, *55*, 402–417.
- (12) Medford, A. J.; Wellendorff, J.; Vojvodic, A.; Studt, F.; Pedersen, F. A.; Jacobsen, K. W.; Bligaard, T.; Nørskov, J. K. Assessing the Reliability of Calculated Catalytic Ammonia Synthesis Rates. *Science* **2014**, *345*, 197–200.
- (13) Wellendorff, J.; Lundgaard, K. T.; Jacobsen, K. W.; Bligaard, T. mBEEF: An Accurate Semi-Local Bayesian Error Estimation Density Functional. *J. Chem. Phys.* **2014**, *140*, 144107.
- (14) Wellendorff, J.; Lundgaard, K. T.; Mogelhoff, A.; Petzold, V.; Landis, D. D.; Nørskov, J. K.; Bligaard, T.; Jacobsen, K. W. Jacobsen, Density Functionals for Surface Science: Exchange-Correlation Model Development with Bayesian Error Estimation. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2012**, *85*, 235149.
- (15) Kulik, H. J.; Cococcioni, M.; Scherlis, D. A.; Marzari, N. Density Functional Theory in Transition-Metal Chemistry: A Self-Consistent Hubbard U Approach. *Phys. Rev. Lett.* **2006**, *97*, 103001.
- (16) Kulik, H. J. Perspective: Treating Electron Over-Localization with the DFT+U Method. *J. Chem. Phys.* **2015**, *142*, 240901.
- (17) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (18) Hammer, B.; Hansen, L. B.; Nørskov, J. K. Improved Adsorption Energetics within Density-Functional Theory using revised Perdew-Burke-Ernzerhof Functionals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1999**, *59*, 7413–7421.
- (19) Zhang, Y.; Yang, W. Comment on “Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1998**, *80*, 890.
- (20) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid Functionals Based on a Screened Coulomb Potential. *J. Chem. Phys.* **2003**, *118*, 8207.
- (21) Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125*, 194101.
- (22) Rencher, A. C. *Methods of Multivariate Analysis*; John Wiley and Sons, Inc.: Hoboken, NJ, 2007.
- (23) Plessis, S.; Carrasco, N.; Pernot, P. Knowledge-Based Probabilistic Representations of Branching Ratios in Chemical Networks: The Case of Dissociative Recombinations. *J. Chem. Phys.* **2010**, *133*, 134110.
- (24) Ammal, S. C.; Heyden, A. Origin of the Unique Activity of Pt/TiO₂ Catalysts for the Water Gas Shift Reaction. *J. Catal.* **2013**, *306*, 78–90.
- (25) Oettel, C.; Rihko-Struckmann, L.; Sundmacher, K. Characterisation of the Electrochemical Water Gas Shift Reactor (EWGSR) Operated with Hydrogen and Carbon Monoxide Rich Feed Gas. *Int. J. Hydrogen Energy* **2012**, *37*, 11759–11771.
- (26) Haryanto, A.; Fernando, S.; Murali, N.; Adhikari, S. Current Status of Hydrogen Production Techniques by Steam Reforming of Ethanol: A Review. *Energy Fuels* **2005**, *19*, 2098–2106.
- (27) Holladay, J. D.; Hu, J.; King, D. L.; Wang, Y. An Overview of Hydrogen Production Technologies. *Catal. Today* **2009**, *139*, 244–260.
- (28) Cortright, R. D.; Davda, R. R.; Dumesic, J. A. Hydrogen from Catalytic Reforming of Biomass-Derived Hydrocarbons in Liquid Water. *Nature* **2002**, *418*, 964–967.
- (29) Song, C. S. Fuel Processing for Low-Temperature and High-Temperature Fuel Cells: Challenges, and Opportunities for Sustainable Development in the 21st Century. *Catal. Today* **2002**, *77*, 17–49.
- (30) Trimm, D. L. Minimisation of Carbon Monoxide in a Hydrogen Stream for Fuel Cell Application. *Appl. Catal., A* **2005**, *296*, 1–11.
- (31) Bartholomew, C. H.; Farrauto, R. J. *Fundamentals of Industrial Catalytic Processes*; Wiley: Hoboken, NJ, 2006.
- (32) Liu, Y.; Fu, Q.; Flytzani-Stephanopoulos, M. Preferential Oxidation of CO in H₂ over CuO-CeO₂ Catalysts. *Catal. Today* **2004**, *93–95*, 241–246.
- (33) Suh, D. J.; Kwak, C.; Kim, J. H.; Kwon, S. M.; Park, T. J. Removal of Carbon Monoxide from Hydrogen-Rich Fuels by Selective Low-Temperature Oxidation over Base Metal Added Platinum Catalysts. *J. Power Sources* **2005**, *142*, 70–74.
- (34) Haruta, M.; Yamada, N.; Kobayashi, T.; Iijima, S. Gold Catalysts Prepared by Coprecipitation for Low-Temperature Oxidation of Hydrogen and of Carbon Monoxide. *J. Catal.* **1989**, *115*, 301–309.
- (35) Bond, G. C.; Thompson, D. T. Catalysis by Gold. *Catal. Rev.: Sci. Eng.* **1999**, *41*, 319–388.
- (36) Haruta, M. Gold as a Novel Catalyst in the 21st Century: Preparation, Working Mechanism and Applications. *Gold Bull.* **2004**, *37*, 27–36.
- (37) Thion, O.; Rachedi, K.; Diehl, F.; Avenier, P.; Schuurman, Y. Kinetics and Mechanism of the Water–Gas Shift Reaction Over Platinum Supported Catalysts. *Top. Catal.* **2009**, *52*, 1940–1945.
- (38) Azzam, K. G.; Babich, I. V.; Seshan, K.; Lefferts, L. A Bifunctional Catalyst for the Single-Stage Water–Gas Shift Reaction in Fuel Cell Applications. Part 2. Roles of the Support and Promoter on Catalyst Activity and Stability. *J. Catal.* **2007**, *251*, 163–171.
- (39) Kalamaras, C. M.; Gonzalez, I. D.; Navarro, R. M.; Fierro, J. L. G.; Efstathiou, A. M. Effects of Reaction Temperature and Support Composition on the Mechanism of Water–Gas Shift Reaction over Supported-Pt Catalysts. *J. Phys. Chem. C* **2011**, *115*, 11595–11610.
- (40) Williams, W. D.; Shekhar, M.; Lee, W. S.; Kispersky, V.; Delgass, W. N.; Ribeiro, F. H.; Kim, S. M.; Stach, E. A.; Miller, J. T.; Allard, L. F. Metallic Corner Atoms in Gold Clusters Supported on Rutile Are the Dominant Active Site during Water–Gas Shift Catalysis. *J. Am. Chem. Soc.* **2010**, *132*, 14018–14020.
- (41) Rodriguez, J. A.; Evans, J.; Graciani, J.; Park, J. B.; Liu, P.; Hrbek, J.; Sanz, J. F. High Water–Gas Shift Activity in TiO₂(110) Supported Cu and Au Nanoparticles: Role of the Oxide and Metal Particle Size. *J. Phys. Chem. C* **2009**, *113*, 7364–7370.
- (42) Vignatti, C.; Avila, M. S.; Apesteguia, C. R.; Garetto, T. F. Catalytic and DRIFTS Study of the WGS Reaction on Pt-Based Catalysts. *Int. J. Hydrogen Energy* **2010**, *35*, 7302–7312.
- (43) Azzam, K. G.; Babich, I. V.; Seshan, K.; Lefferts, L. Bifunctional Catalysts for Single-Stage Water–Gas Shift Reaction in Fuel Cell Applications.: Part 1. Effect of the Support on the Reaction Sequence. *J. Catal.* **2007**, *251*, 153–162.
- (44) Kalamaras, C. M.; Panagiotopoulou, P.; Kondarides, D. I.; Efstathiou, A. M. Kinetic and Mechanistic Studies of the Water–Gas Shift Reaction on Pt/TiO₂ catalyst. *J. Catal.* **2009**, *264*, 117–129.
- (45) Green, I. X.; Tang, W. J.; Neurock, M.; Yates, J. T. Spectroscopic Observation of Dual Catalytic Sites During Oxidation of CO on a Au/TiO₂ Catalyst. *Science* **2011**, *333*, 736–739.
- (46) Rashkeev, S. N.; Lupini, A. R.; Overbury, S. H.; Pennycuik, S. J.; Pantelides, S. T. Role of the Nanoscale in Catalytic CO Oxidation by Supported Au and Pt Nanostructures. *Phys. Rev. B: Condens. Matter Phys.* **2007**, *76*, 035438.
- (47) Pazmino, J. H.; Shekhar, M.; Williams, W. D.; Akatay, M. C.; Miller, J. T.; Delgass, W. N.; Ribeiro, F. H. Metallic Pt as Active Sites for the Water–Gas Shift Reaction on Alkali Promoted Supported Catalysts. *J. Catal.* **2012**, *286*, 279–2861.
- (48) Ding, K.; Gulec, A.; Johnson, A. M.; Schweitzer, N. M.; Stucky, G. D.; Marks, L. D.; Stair, P. C. Identification of Active Sites in CO

Oxidation and Water-Gas Shift over Supported Pt Catalysts. *Science* **2015**, 350, 189–192.

(49) Fu, Q.; Saltsburg, H.; Stephanopoulos, M. F. Active Nonmetallic Au and Pt Species on Ceria-Based Water-Gas Shift Catalysts. *Science* **2003**, 301, 935–938.

(50) Yang, M.; Liu, J.; Lee, S.; Zugic, B.; Huang, J.; Allard, L. F.; Flytzani-Stephanopoulos, M. A Common Single-Site Pt(II)–O(OH)-x- Species Stabilized by Sodium on “Active” and “Inert” Supports Catalyzes the Water-Gas Shift Reaction. *J. Am. Chem. Soc.* **2015**, 137, 3470–3473.

(51) Ammal, S. C.; Heyden, A. Modeling the Noble Metal/TiO₂ metal/TiO₂(110) Interface with Hybrid DFT Functionals: A Periodic Electrostatic Embedded Cluster Model Study. *J. Chem. Phys.* **2010**, 133, 164703.

(52) Ammal, S. C.; Heyden, A. Water–Gas Shift Catalysis at Corner Atoms of Pt Clusters in Contact with a TiO₂ (110) Support Surface. *ACS Catal.* **2014**, 4, 3654–3662.

(53) Rubin, D. B.; Thayer, D. T. EM Algorithms for ML Factor Analysis. *Psychometrika* **1982**, 47, 69–76.

(54) Terejanu, G.; Upadhyay, R. R.; Miki, K. Bayesian Experimental Design for the Active Nitridation of Graphite by Atomic Nitrogen. *Exp. Therm. Fluid Sci.* **2012**, 36, 178–193.

(55) Kullback, S.; Leibler, R. A. On Information and Sufficiency. *Ann. Math. Stat.* **1951**, 22, 79–86.

(56) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, 140, A1133.

(57) Adamo, C.; Scuseria, G. E.; Barone, V. Accurate Excitation Energies from Time-Dependent Density Functional Theory: Assessing the PBE0Model. *J. Chem. Phys.* **1999**, 111, 2889.

(58) Xu, Z.; Joshi, Y. V.; Raman, S.; Kitchin, J. R. Accurate Electronic and Chemical Properties of 3d Transition Metal Oxides Using a Calculated Linear Response U and a DFT + U(V) Method. *J. Chem. Phys.* **2015**, 142, 144701.

(59) Henkelman, G.; Uberuaga, B. P.; Jonsson, H. A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *J. Chem. Phys.* **2000**, 113, 9901.

(60) Heyden, A.; Bell, A. T.; Keil, F. J. Efficient Methods for Finding Transition States in Chemical Reactions: Comparison of Improved Dimer Method and Partitioned Rational Function Optimization Method. *J. Chem. Phys.* **2005**, 123, 224101.

(61) Iida, H.; Igarashi, A. Characterization of a Pt/TiO₂ (Rutile) Catalyst for Water Gas Shift Reaction at Low-Temperature. *Appl. Catal., A* **2006**, 298, 152–160.

(62) Panagiotopoulou, P.; Christodoulakis, A.; Kondarides, D. I.; Boghosian, S. Particle Size Effects on the Reducibility of Titanium Dioxide and its Relation to the Water–Gas Shift Activity of Pt/TiO₂ Catalysts. *J. Catal.* **2006**, 240, 114–125.

(63) We plan to make our code publically available during summer 2016.