Acceleration of Catalyst Discovery with Easy, Fast, and Reproducible Computational Alchemy

Charles Griego[1], John R. Kitchin[2], and John A. Keith[1]

[1]University of Pittsburgh
[2]Carnegie Mellon University (CMU)

June 12, 2020

Abstract

The expense of quantum chemistry calculations significantly hinders the search for novel catalysts. Here, we provide a tutorial for using an easy and highly cost-efficient calculation scheme called alchemical perturbation density functional theory (APDFT) for rapid predictions of binding energies of reaction intermediates and reaction barrier heights based on Kohn-Sham density functional theory reference data. We outline standard procedures used in computational catalysis applications, explain how computational alchemy calculations can be carried out for those applications, and then present bench marking studies of binding energy and barrier height predictions. Using a single OH binding energy on the Pt(111) surface as a reference case, we use computational alchemy to predict binding energies of 32 variations of this system with a mean unsigned error of less than 0.05 eV relative to single-point DFT calculations. Using a single nudged elastic band calculation for CH_4 dehydrogenation on Pt(111) as a reference case, we generate 32 new pathways with barrier heights having mean unsigned errors of less than 0.3 eV relative to single-point DFT calculations. Notably, this easy APDFT scheme brings no appreciable computational cost once reference calculations are done, and this shows that simple applications of computational alchemy can significantly impact DFT-driven explorations for catalysts. To accelerate computational catalysis discovery and ensure computational reproducibility, we also include Python modules that allow users to perform their own computational alchemy calculations.

Keywords — Computational catalysis, density functional theory (DFT), adsorption energies, nudged elastic band calculations, binding energies, barrier heights

Introduction

Advances in computational chemistry open new possibilities for impressively large scale computational screening of hypothetical catalysts across materials space.[1],[2],[3] However, productively leveraging high throughput screening has been challenging. For useful and insightful predictions, computational screening studies must be reproducible while also i) determining important active sites that are stable under specified environmental conditions on large numbers of material compositions and ii) elucidating important elementary reaction steps with barrier heights that are needed for kinetic modeling. To date, most computational catalysis studies can address one of these points at a time, but new advances are needed to make it possible to address both points at the same time.

Reliably accurate quantum chemistry (QC) methods, e.g. Kohn-Sham Density Functional Theory (DFT), continue to play a central role for computational catalysis studies. Even though the computational cost of QC calculations continues to decrease to allow even more calculations possible in the future, one should not expect that the two points above can be addressed at the same time using QC calculations without invoking severe approximations. For example, the cost of most DFT calculations limits their utility in screening studies to $O(10^3$-$10^4)$ of kinds of active sites. If one considers just binary pairs of twenty different metals, there are 190 possible binary alloys. Each alloy in turn has an enormous configuration space, e.g. for a 55-atom nanoparticle, there are close to 2^{55} possible configurations that potentially need to be considered.

Thus, we need a computationally inexpensive but sufficiently accurate approach to estimate which alloy configurations are important enough to warrant further study with QC calculations.

A similar challenge exists with calculating reaction barriers. These computations are much more expensive than stability or adsorption energy calculations. Only the most relevant barriers ideally would be studied with costly QC methods, but determining whether a barrier is relevant requires iterative kinetic analyses that themselves require numerous barrier height predictions to be useful. The key point is these studies would be significantly accelerated if barriers could be estimated quickly and accurately enough to determine if more QC calculations were needed.

The present tutorial shows how to apply an approximate method that has promise to accelerate computational catalysis screening studies. The method has previously been called "computational alchemy",[4],[5],[6],[7] but more recently termed "alchemical perturbation density functional theory"[8] to better distinguish it to other forms of "computational alchemy".[9] The tutorial begins with an overview of how computational catalysis is normally done, we then provide an introduction to the theory of computational alchemy, and then we benchmark a set of adsorbate binding energies (BE) and activation energies (E_a). Large amounts of computational alchemy data can be obtained from a single set of DFT calculations by using updatable and reproducible procedures embedded within Jupyter notebooks. This is expected to make it useful for accelerating computational catalysis screening predictions of adsorption energies and reaction barriers.

Conventional Computational Catalysis

Catalytic reaction mechanisms depend on the interplay of reaction thermodynamics and kinetics. The present work assumes that these can often be modeled quite well with standard DFT calculations, but bringing this level of accuracy already brings substantial computational cost. More complicated systems will be expected to require higher level QC calculations that bring even greater costs.

The thermodynamics of elementary reaction steps can be assessed by computing reaction intermediate BEs. Each BE determination normally requires three separate QC electronic energy calculations, each preferably modeled using geometrically relaxed structures, and suitably accurate levels of theory. Appropriate zero-point energy, thermal, and entropy corrections can be included as well, but these are neglected here for simplicity. The required calculations include the model of the catalyst surface without the adsorbate (site), a model of the adsorbate not interacting with the catalyst (ads), and a model for the catalyst surface with the adsorbate bound (ads-site). A BE is then calculated with the following equation:

$$BE = E_\text{site} + E_\text{ads} - E_\text{ads-site} \tag{1}$$

In this convention, positive binding energies indicate thermodynamically favorable adsorptions. This equation clearly highlights a problem with high throughput screening studies that rely on QC calculations. Any BE calculation for a single adsorbate on any hypothetical surface site usually requires two separate QC calculations (for $E_\text{ads-site}$ and E_site), and those might take anywhere from minutes to weeks to complete on modern super computers. Predictive BEs also may require more considerations of multiple adsorbate configurations as well as adsorbate-induced surface reconstruction and segregation.

Relatively efficient DFT calculations for N electron systems typically scale as N^3 or greater, making these calculations quite expensive. This unfavorable scaling is especially significant when calculating E_a. The Nudged Elastic Band (NEB) algorithm,[10] which is now a standard approach for predicting E_a, works by computing a series of constrained geometric 'images' that are eventually optimized to a minimum energy pathway along a potential energy surface. Since standard NEB calculations generally require the optimization of $10-20$ images per pathway, they bring significant computational cost that limits applications in high-throughput screening. A conventional approach to address this is with linear scaling relations,[11],[12],[13] but an open question is whether alternative approaches might be more accurate and useful.

2

APDFT

The poor scaling of QC calculations in computational catalysis applications motivate the need for developing faster methods to estimate BEs and E_as. Brute-force computations driven exclusively by DFT calculations are not expected to be aproductive for large-scale explorations for hypothetical catalysts. Machine learning models for catalysis prediction are becoming a more common alternative,[14],[15],[3] but a downside of most of these approaches is that they need a lot of data for training, they can unreliably extrapolate to new systems, and they are often harder to intuitively or physically interpret. An alternative and general approach being championed by von Lilienfeld and co-workers [4],[5],[16] is to use gradient-supported methods that consider how a change to a material results in a change in a specific property without the explicit QC calculation of the property itself. These approaches are essentially a special class of quantitative structure property relationship (QSPR) methods,[17] but while most QSPR models are empirically fit, the end goal here is to make computationally efficient predictions based on first principles-derived entities such as a wavefunction, an electron density, or an electrostatic potential.

We now provide a tutorial on how a simple form of APDFT can be employed for computational catalysis applications [18],[19],[7],[20]. To begin, we consider the BE calculation for an OH molecule on Pt(111). This system will be referred as our reference state and subsequently labeled with $\lambda = 0$. Since zero-point, thermal, and entropy corrections will be neglected out of simplicity, its BE will be labeled as $\Delta E^0|_{\lambda=0}$. We now will consider a new system where the Pt(111) surface has been doped with a new element that results in a new binding site. This new state will be labeled as $\lambda = 1$, and its BE will be labeled as $\Delta E^0|_{\lambda=1}$. APDFT can be used to predict $\Delta E^0|_{\lambda=1}$ by relating it to $\Delta E^0|_{\lambda=0}$ using a thermodynamic cycle shown in Fig. 1.
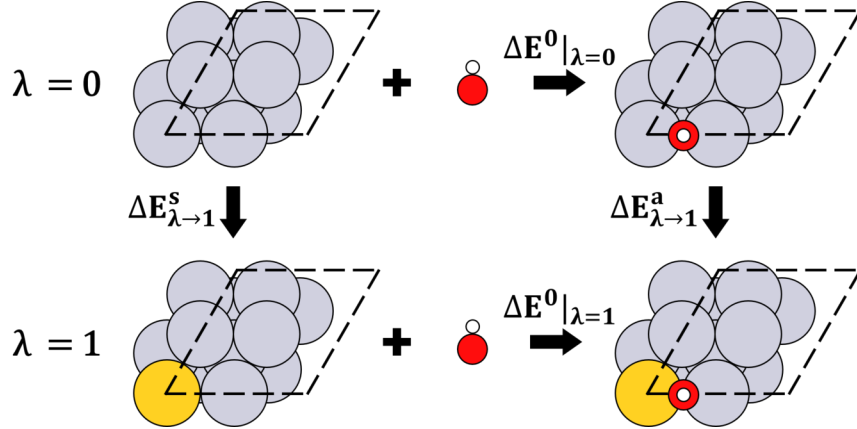

Figure 1: Thermodynamic cycle depicting the binding energies (BEs) of an adsorbate on a surface (horizontal legs) and atomic transmutations (vertical legs). $\Delta E^0|_{\lambda=0}$ and $\Delta E^0|_{\lambda=1}$ denote the BEs for the top and bottom horizontal legs, respectively. $\Delta E^s_{\lambda\to1}$ and $\Delta E^a_{\lambda\to1}$ denote the energy change associated with the atomic transmutation for the left (s = surface) and right (a = ads-site) vertical legs, respectively.

According to the cycle, we have Eq. 2:

$$\Delta E^0|_{\lambda=0} + \Delta E^a_{\lambda\to1} = \Delta E^0|_{\lambda=1} + \Delta E^s_{\lambda\to1} \tag{2}$$

Where $\Delta E^0|_{\lambda=0}$ will be obtained using Eq. 1, $\Delta E^s_{\lambda\to1}$ is the energy change of the bare catalyst surface (s = site) when doped with a new element, and $\Delta E^a_{\lambda\to1}$ is the same type of change for the system of the catalyst site with a bound adsorbate (a = ads-site). Upon rearrangement of Eq. 2, the change in BE due to the transition from the $\lambda = 0$ to the $\lambda = 1$ states is the difference of these two unknown terms:

$$\Delta BE = \Delta E^0|_{\lambda=1} - \Delta E^0|_{\lambda=0} = \Delta E^a_{\lambda\to1} - \Delta E^s_{\lambda\to1} \tag{3}$$

3

We now will demonstrate how to calculate the right side of Eq. 3 using APDFT to obtain ΔBE. APDFT relates $\Delta E^0|_{\lambda=1}$ to $\Delta E^0|_{\lambda=0}$ with perturbation theory, specifically by approximating the exact result as a Taylor series expansion with the thermodynamic state function (ΔE^0), Eq. 4.[19]

$$\Delta E^0|_{\lambda=1} = \Delta E^0|_{\lambda=0} + \partial_\lambda \Delta E^0 \Delta\lambda + \frac{1}{2}\partial_\lambda^2 \Delta E^0 \Delta\lambda^2 + ... \tag{4}$$

Here, the predicted BE $(\Delta E^0|_{\lambda=1})$ is approximated as the reference BE $(\Delta E^0|_{\lambda=0})$ plus additional corrections from the Taylor series based on so-called "alchemical derivatives", which are other terms resulting from the transformation from the $\lambda = 0$ state to the $\lambda = 1$ state. The *nth* alchemical derivative with respect to λ is denoted by $\partial_\lambda^n \Delta E^0$. For the easiest applications of APDFT, one can simply truncate the expansion in Eq. 4 to first order and assign $\Delta\lambda = 1$, but higher order corrections for non-periodic computational schemes have been developed by von Lilienfeld and co-workers.[21] The first-order approximation for APDFT is shown by Eq. 5.

$$\partial_\lambda \Delta E^0 = \sum_I \Delta\mu_{nI} \partial_\lambda N_I - \sum_I \Delta F_I \partial_\lambda R_I + \Delta\mu_e \partial_\lambda N_e \tag{5}$$

Here, the first term accounts for the nuclear chemical potential gradient $(\Delta\mu_{nI})$ due to variation in nuclear charge (N_I) from the $\lambda = 0$ to the $\lambda = 1$ state. The second term accounts for energy gradients due to the forces on atoms (ΔF_I) resulting from changes in atomic positions (R_I) from the $\lambda = 0$ to the $\lambda = 1$ state. The third term accounts for the electronic chemical potential gradient $(\Delta\mu_e)$ due to variation in total number of electrons (N_e) from the $\lambda = 0$ to the $\lambda = 1$ state.

Two additional approximations are typically employed with simple applications of first order APDFT used for computational catalysis studies of extended surfaces. The first approximation is to assume that nuclear positions for the $\lambda = 0$ state and the $\lambda = 1$ state are the same. While changing a material's composition will certainly impact interatomic forces and result in relaxations to achieve a minimum energy state, it will be assumed for simplicity that the relaxation energy contributions due to the doping atom in the $\lambda = 1$ states for the "ads-site" and "site" calculations are similar and thus largely cancel in the thermodynamic cycle represented in Fig. 1. Mathematically, the result is that $\delta_\lambda R_I$ in the second term of Eq. 5 becomes zero, and thus the second term can be neglected. The second approximation is to ensure that the $\lambda = 0$ state and the $\lambda = 1$ state have the same number of electrons so that $\delta_\lambda N_e$ becomes zero so that the third term becomes zero. (Note that this is not a general requirement for APDFT, but it is a practical way to ensure that a total surface change density under periodic boundary conditions remains reasonably physical in these specific applications.) To apply this constraint, one must use "isoelectronic transmutations" so that the $\lambda = 1$ system has the same number of electrons as the $\lambda = 0$ system. For example, in a Pt(111) slab, if one Pt atom is transmuted into an Au atom $(\Delta Z = +1)$, a second Pt atom must be transmuted into an Ir atom $(\Delta Z = -1)$. This results in a zero net change in atomic number for the new slab $(\sum_{n=1}^{\text{no. of atoms}} \Delta Z = 0)$. While the second transmutation is a departure from the system of interest, it can be made to occur far from the site of interest so that it has no significant effect on the BE of that site.

To summarize, the simplest application of first order APDFT for predicting a BE for a hypothetical catalyst surface will consider only fixed atomic coordinates and only net isoelectronic transmutations, and this simplifies Eq. 5 into Eq. 6.

$$\partial_\lambda \Delta E^0 = \sum_I \Delta\mu_{nI} \partial_\lambda N_I \tag{6}$$

The nuclear chemical potential, $\Delta\mu_{nI}$ (also called the alchemical potential), is defined as the change in atomic electrostatic potential between the "site" and the "ads-site" reference states. This is represented as an array of differences in electrostatic potentials, and so the elements of this array have units of energy/charge. Note that this procedure accounts for changes in electrostatic potentials of atoms on the catalyst surface after relaxations due to molecular adsorption have taken place. This term mathematically refers to the difference shown by Eq. 7.

$$\Delta\mu_{nI} = \mu_n^{\text{ads-site}}(R_I^{\text{ads-site}}) - \mu_n^{\text{site}}(R_I^{\text{site}}) \tag{7}$$

These atom-centered electrostatic potentials for the "ads-site" and "site" reference states can be obtained with VASP,[22] a widely-used DFT code. Fig. 2a depicts how alchemical potentials are constructed as an array of differences in electrostatic potentials for atoms from both the "site" and "ads-site" states. The nuclear charge variation quantity, $\partial_\lambda N_I$, in Eq. 6 now should be accounted for. When one transmutes atoms from a $\lambda = 0$ state to form a hypothetical $\lambda = 1$ state, one notes changes in nuclear charges. Specifically, one uses an array of differences in nuclear charges due to the formation of the $\lambda = 1$ state, and so the elements in this array have units of charge. Fig. 2b depicts how nuclear charge variations are constructed as an array of differences in atomic numbers of the atoms in the $\lambda = 0$ state and the $\lambda = 1$ state. Note that this array will be the same whether it is constructed for a "site" or "ads-site" state since they both will correspond to the same $\lambda = 1$ state. The overall calculation of the alchemical derivative used to obtain ΔBE between the $\lambda = 0$ state and the $\lambda = 1$ state is simply a dot product of the two arrays given above, and this gives a scalar that has units of energy (Fig. 2c).
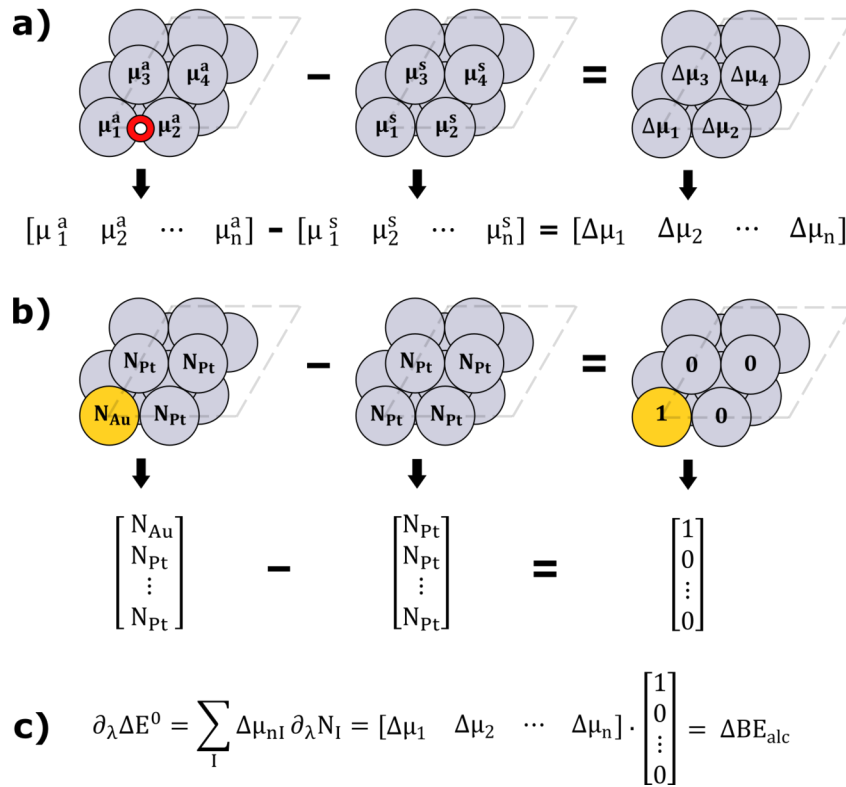

Figure 2: Illustrations of steps to compute corrections to binding energies (BEs) based on APDFT. **a)** Constructing the alchemical potential array with differences between μ_i^a and μ_i^s, the electrostatic potentials for the ith atom in the ads-site (a) and site (s) systems, respectively. The adsorbate atoms in the ads-site state are conventionally not included in the array because they are not subject to transmutations. **b)** Constructing the nuclear charge variation array with differences in atomic number for each atom in the system before and after the transmutation. **c)** Illustrations of Eqs. 6 and 11 (see below) as a dot product of the two arrays constructed in **a)** and **b)**. (Note that the illustrations in **a)** and **b)** do not show an isoelectronic change and the result of the equality shown in **c)** is provided to show a simple visual representation.)

For additional clarity, we can return to the two unknown terms ($\Delta E_{\lambda \to 1}^a$ and $\Delta E_{\lambda \to 1}^s$) in Eq. 3 that compute ΔBE according to the cycle in Fig. 1. We now show the direct relation to the results of Eqs. 6 and 7 and

5

the steps in Fig. 2. The left leg of Fig. 1 and the expression in Fig. 2b both illustrate a reference Pt(111) surface "site" undergoing an alchemical transmutation of one Pt atom into an Au atom. Using Eq. 6, a first-order approximation of the energy change is:

$$\Delta E^s_{\lambda \to 1} = \sum_I \mu^{\text{site}}_{nI}(R^{\text{site}}_I)\partial_\lambda N^{\text{site}}_I \tag{8}$$

Note that the illustrations in Fig. 1 and Fig. 2b do not show an isoelectronic change to provide a simple visual representation, but the arrays used in Eq. 8 should account for them. The energy change for the same alchemical transmutation done to the "ads-site" species is:

$$\Delta E^a_{\lambda \to 1} = \sum_I \mu^{\text{ads-site}}_{nI}(R^{\text{ads-site}}_I)\partial_\lambda N^{\text{ads-site}}_I \tag{9}$$

For the cycle in Fig. 1 to be true, the transmutations done to the "slab" and "ads-site" states must be exact, and this results in $\partial_\lambda N^{\text{site}}_I = \partial_\lambda N^{\text{ads-site}}_I = \partial_\lambda N_I$. Returning to Eq. 3, and using Eqs. 6 through 9:

$$\Delta E^a_{\lambda \to 1} - \Delta E^s_{\lambda \to 1} = \sum_I [\mu^{\text{ads-site}}_{nI}(R^{\text{ads-site}}_I) - \mu^{\text{site}}_{nI}(R^{\text{site}}_I)]\partial_\lambda N_I = \sum_I \Delta\mu_{nI}\partial_\lambda N_I \tag{10}$$

Finally, we combine Eqs. 3 and 10 to rewrite our simple approximation, illustrated in Fig. 2c, for the predicted BE change using computational alchemy:

$$\Delta BE_{alc} = \sum_I \Delta\mu_{nI}\partial_\lambda N_I \tag{11}$$

It is interesting that BEs from first order APDFT appear to be valid for cases that are and are not primarily arising from d-orbital contributions. In contrast, the well-established "d-band model" for predicting adsorbate BEs on transition metals has known shortcomings when d-orbital contributions to binding are less significant.[23] One known limitation for BE predictions from first order APDFT predictions is that they are less accurate when alchemical derivatives become large (e.g. when a system undergoes multiple and/or large transmutations where ΔZ is larger than 1).[7] Another limitation is that these simple APDFT treatments appear not to be valid for materials that have a band gap.[20] While more is being learned about the promise and limitations, we see APDFT as a potentially transformative model for applied computational chemistry communities and especially for computational catalysis. We will now show benchmarking calculations to validate the use of computational alchemy for predictions of BE and E_a in two different examples.

Results and Discussion

Benchmarking adsorbate BEs

We first benchmark BEs for OH adsorption on a four-layer, 2×2 Pt(111) surface model that contains four Pt atoms in each layer. For the binding configuration of OH on the surface, the fourth layer has two unique atoms that are transmutable. Thus, we created hypothetical materials by transmuting one of the eight atoms in the top two layers by $\Delta Z = \pm 1$ to convert a Pt near a binding site into Au or Ir and one of the two atoms in the fourth layer by $\Delta Z = \mp 1$. This generated 32 unique cases that computational alchemy can be used for predictions based on a single DFT binding energy calculation.
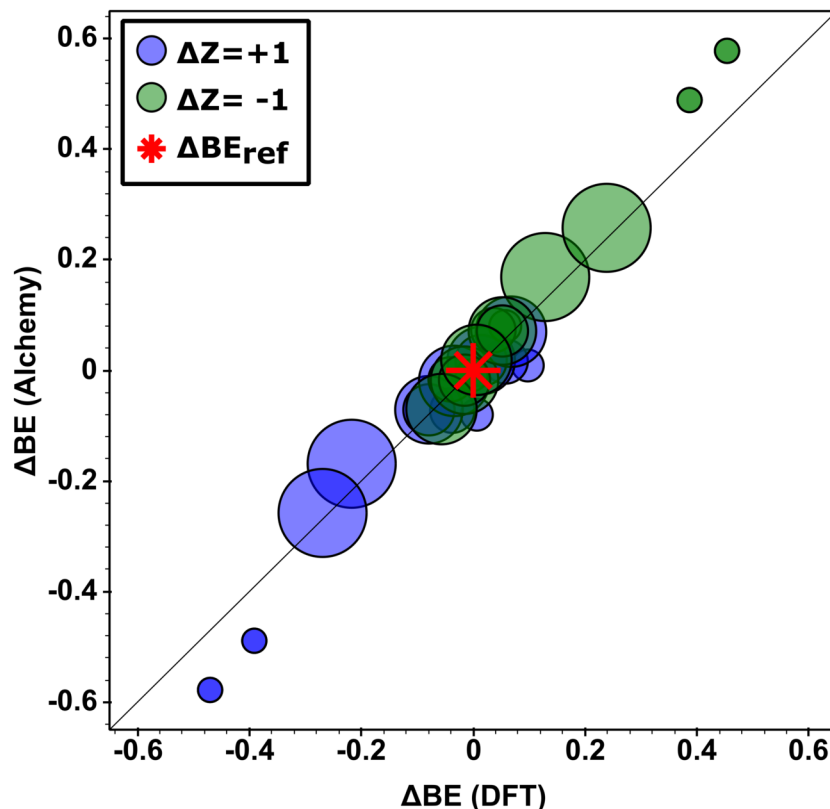
Figure 3: ΔBE predictions for OH adsorption on 32 hypothetical alloys of Pt made by $\Delta Z = \pm 1$ transmutations to the reference slab. Energy differences are relative to the reference calculation and reported in eV. The size of the data points corresponds to the distance of the transmutation site from the adsorbate (large points are transmutations far from the adsorption site).

Fig. 3 shows a benchmarking comparison of first order APDFT calculations compared to corresponding single point DFT energy calculations. Note that these data were previously reported elsewhere,[7] but the data shown now depict the relative location of transmuted atoms in the slab model to the OH adsorbate. Larger data points depict cases where transmuted atoms are further from the adsorbate while smaller data points depict cases where they are closer. These predictions have a mean absolute error of 0.045 eV. The first order APDFT is notably less accurate for transmutations made directly at the binding site, where the alchemical derivatives are the greatest.[7]

In summary, Fig. 3 reiterates that first order APDFT can quite accurately predict how BEs changes on hypothetical alloys with simple algebraic computations based on a single set of QC calculations and corresponding electrostatic potentials. This provides a computational lever to increase the utility of a single BE calculation by enabling the estimation of (in this case) about 30 additional BEs. Even more BEs can be predicted by employing more and/or larger (e.g. $\Delta Z = \pm 2$) transmutations, but we have found these will result in larger errors that would need to be remedied for accuracy. Not only does this tool give users BE predictions, it also gives users physical insight into which nearby sites are the most important for adsorbate binding. In principle, this screening allows us to eliminate cases that result in an insignificant change in ΔBE, and thus allow more careful attention to cases that result in more significant changes in BEs using more refined QC methods and/or experimental studies.

7

Benchmarking reaction barriers

To benchmark first order APDFT predictions of reaction barriers, we use CH_4* dehydrogenation on Pt(111) as a reference system (Eq. 12), where each species bound to a surface site is denoted by *:

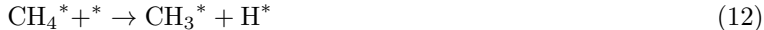$$CH_4{}^* + {}^* \rightarrow CH_3{}^* + H^*$$

(12)

Fig. 4 shows snapshots of the NEB reaction pathway for this process. The reactant state and first image in the NEB (a) contains CH_4* geometrically relaxed to a distance of 3.7 Å from an ontop site. The reaction proceeds as one C-H bond breaks and a hydrogen is adsorbed to an adjacent ontop site. In the transition state, the detaching hydrogen is drawn toward a bridge site. The product state contains both a CH_3* and an H* bound at 2.1 and 1.5 Å at ontop sites, respectively. Other work reports that H* also binds favorably to an fcc site in the product state.[24] However, our system is among many reasonable choices of reference points to predict changing trends of various possible reaction pathways on many hypothetical catalysts.
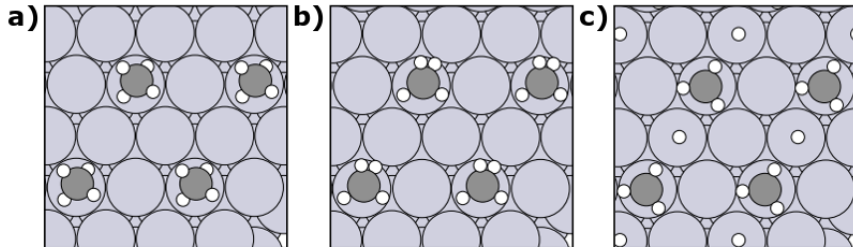

Figure 4: Snapshots of the (a) reactant, (b) transition, and (c) product states of the dehydrogenation of CH_4 adsorbed at $\frac{1}{4}$ coverage of ontop sites on Pt.

To compute barrier heights using first order APDFT, we apply the BE procedure from above on each of the images from the NEB calculation. From this, we can generate a variety of different pathways for up to 32 different hypothetical alloy configurations as shown in Fig. 5. Fig. 5a shows predicted reaction energy profiles following Eq. 12. The reference energy profile calculated using DFT is denoted by red asterisks, while the solid blue line denotes the most affected Au alloys and the solid green line denotes the most affected Ir alloys. The other lines pertain to all other cases where the alloy systems result in a negligible change relative to the reference system (due to very small alchemical derivatives present in these cases). Note that energy profiles for similar systems overlap and may not be distinguishable. Figure 5b shows a parity plot for ΔE_a values relative to the reference calculation. There is generally an increase in barrier when the transmuted site becomes Au and a decrease when the site becomes Ir. Notably, the largest errors in barrier heights are 0.3 eV relative to DFT benchmarks (for alloys that exhibited the highest alchemical derivatives), but the vast majority of other data are more accurate. Even though first order APDFT can exhibit errors as large as 0.3 eV in barrier heights for reference system doped with just a single atom, it is promising that such simple approximations can be useful to calculate a computationally expensive descriptor that guides screening studies. Furthermore, as the source for errors in different APDFT approximations become better understood, it becomes reasonable to imagine that more accurate approaches might be developed based on APDFT that would have a predictive power comparable (or even indistinguishable) to standard DFT.
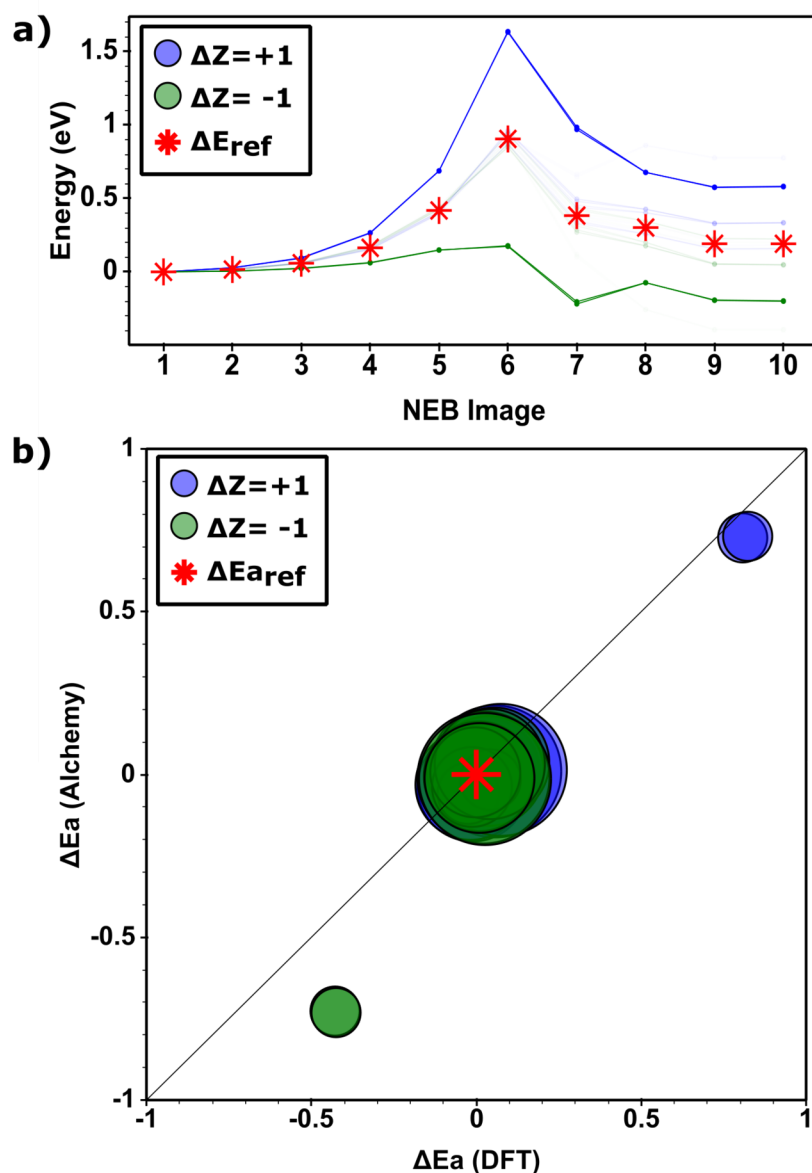
Figure 5: **a)** Energy profiles for the CH_4 dehydrogenation mechanism on hypothetical alloys of Pt. The reference pathway occurs on pure Pt and is denoted with red asterisks. The most significant effect from a transmutation with $\Delta Z = +1$ is shown in blue, the most significant effect from a transmutation with $\Delta Z = -1$ is shown in green, and other reaction pathways computed with alchemy are shown in light blue/green. **b)** First order APDFT benchmarking of the change in ΔE_a for CH_4 dehydrogenation on 32 hypothetical alloys of Pt made by $\Delta Z = \pm 1$ relative to the reference barrier (energies in eV). The size of the data points correspond to the distance of the transmutation site from the adsorbate in the reactant state.

9

Usability and reproducibility with computational alchemy

Python scripts that enable a workflow with first order APDFT suitable for use with VASP are open source and continuously improving. Additionally, we offer hands-on outlines for calculations via Jupyter notebooks so that readers can reproduce our analyses presented above. Interactive forms of Figs. 3 and 5 appear in the following two notebooks:

- Binding Energy Predictions
- Transition State Barrier Height Predictions

Readers may then easily use first order APDFT with their own VASP reference calculations by installing our Python modules and following the procedures written in these notebooks. These modules rely on an installation of the Atomic Simulation Environment package [25] for reading, writing, and visualizing atomic structures, and pristine copies of the POSCAR, CONTCAR, OSZICAR, and OUTCAR files from all VASP calculations are required. Users of these modules will be able to do the following:

- Calculate atom-centered alchemical potentials
- Visualize a heat-map of the alchemical derivatives within their surface model
- Create a custom set of hypothetical materials using isoelectronic transmutations
- Calculate alchemical derivatives for the hypothetical systems, and depending on the type of reference calculation set, compute BE, E_a, or other thermodynamic descriptors

Outlook/Known limitations

APDFT is a promising gradient-supported method for leveraging QC calculations in large scale screening. This method relies on alchemical derivatives which are straight-forward to compute and which can be used for accurate predictions of binding energies and activation energies on electronically conductive systems relevant for computational catalysis. There are still notable limitations that need to be considered before widespread use of computational alchemy would be possible.

A significant challenge with APDFT is that the simplest first order approximations appear to have some material dependencies, and predictions appear to fail when used on systems having band gaps. A possible remedy we found is to add metal dopants into the semiconducting or insulating system. This allows bandgaps to collapse, and computational alchemy appears to be valid.[20] Future work will address the physical descriptions of these systems and their relation to existing models such as the d-band model. We[7] and other groups[6] have also noted that predictions from first order APDFT become less accurate with the magnitude of the alchemical derivatives that are in play. There are possible options to address this ranging from employing higher-order corrections and/or training machine learning algorithms to correct errors for problematic systems.[26]

Conclusions

We have presented a tutorial for how first order APDFT can be used as a tool to dramatically accelerate predictions of computational catalysis descriptors. This is a means to predict the impact of a material change on a calculated energy with essentially no computational cost once reference calculations have been obtained. Here, we demonstrated how to calculate binding energies and activation energies and provided open source scripts for doing so. Using these scripts should enable rapid determinations of which atoms within a catalyst should be changed to impact catalysis. Further improvements to these existing schemes are expected with the use of second order corrections for improved physical descriptions and coupling to machine learning models for improved accuracy.

10

Acknowledgements

References

[1]A. Kulkarni, S. Siahrostami, A. Patel, J. K. Nørskov, *Chem. Rev.* **2018**, *118*, 2302–2312.

[2]J. Greeley, *Annu. Rev. Chem. Biomol. Eng* **2016**, *7*, 605–635.

[3]K. Tran, Z. W. Ulissi, *Nature Catal.* **2018**, *1*, 696–703.

[4]O. A. von Lilienfeld, R. D. Lins, U. Rothlisberger, *Phys. Rev. Lett.* **2005**, *95*, DOI 10.1103/physrevlett.95.153002.

[5]O. A. von Lilienfeld, M. E. Tuckerman, *J. Chem. Phys.* **2006**, *125*, 154104.

[6]M. Muñoz, C. Cárdenas, *Phys. Chem. Chem. Phys.* **2017**, *19*, 16003–16012.

[7]K. Saravanan, J. R. Kitchin, O. A. von Lilienfeld, J. A. Keith, *J. Phys. Chem. Lett.* **2017**, *8*, 5002–5007.

[8]G. F. von Rudorff, O. A. von Lilienfeld, *Physical Chemistry Chemical Physics* **2020**, DOI 10.1039/c9cp06471k.

[9]T. P. Straatsma, J. A. McCammon, *Annual Review of Physical Chemistry* **1992**, *43*, 407–435.

[10]G. Henkelman, B. P. Uberuaga, H. Jónsson, *J. Chem. Phys.* **2000**, *113*, 9901–9904.

[11]F. Abild-Pedersen, J. Greeley, F. Studt, J. Rossmeisl, T. R. Munter, P. G. Moses, E. Skúlason, T. Bligaard, J. K. Nørskov, *Phys. Rev. Lett.* **2007**, *99*, DOI 10.1103/physrevlett.99.016105.

[12]A. Vojvodic, F. Calle-Vallejo, W. Guo, S. Wang, A. Toftelund, F. Studt, J. I. Martínez, J. Shen, I. C. Man, J. Rossmeisl, et al., *J. Chem. Phys.* **2011**, *134*, 244509.

[13]S. Wang, B. Temel, J. Shen, G. Jones, L. C. Grabow, F. Studt, T. Bligaard, F. Abild-Pedersen, C. H. Christensen, J. K. Nørskov, *Catal. Lett.* **2010**, *141*, 370–373.

[14]Z. Li, S. Wang, W. S. Chin, L. E. Achenie, H. Xin, *J. Mater. Chem. A* **2017**, *5*, 24131–24138.

[15]Z. Li, X. Ma, H. Xin, *Catal. Today* **2017**, *280*, 232–238.

[16]M. to Baben, J. O. Achenbach, O. A. von Lilienfeld, *J. Chem. Phys.* **2016**, *144*, 104103.

[17]T. Le, V. C. Epa, F. R. Burden, D. A. Winkler, *Chem. Rev.* **2012**, *112*, 2889–2919.

[18]Y. S. Al-Hamdani, A. Michaelides, O. A. von Lilienfeld, *J. Chem. Phys.* **2017**, *147*, 164113.

[19]D. Sheppard, G. Henkelman, O. A. von Lilienfeld, *J. Chem. Phys.* **2010**, *133*, 084104.

[20]C. D. Griego, K. Saravanan, J. A. Keith, *Adv. Theory Simul.* **2018**, *2*, 1800142.

[21]G. F. von Rudorff, O. A. von Lilienfeld, *J. Phys. Chem. B* **2019**, *123*, 10073–10082.

[22]G. Kresse, J. Furthmüller, *Phys. Rev. B* **1996**, *54*, 11169–11186.

[23]H. Xin, S. Linic, *J. Chem. Phys.* **2010**, *132*, 221101.

[24]R. Zhang, L. Song, Y. Wang, *Appl. Surf. Sci.* **2012**, *258*, 7154–7160.

[25]A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, et al., *J. Phys. Condens. Matter* **2017**, *29*, 273002.

[26]K. Saravanan, Computational High Throughput Searches for Efficient Catalysts - D-Scholarship@Pitt, PhD thesis, http://d-scholarship.pitt.edu/35610/, **2019**.