

# Evaluating quantum alchemy of atoms with thermodynamic cycles: Beyond ground electronic states

Emily A. Eikey,<sup>1</sup> Alex M. Maldonado,<sup>2</sup> Charles D. Griego,<sup>2</sup> Guido Falk von Rudorff,<sup>3</sup> and John A. Keith<sup>2, a)</sup>

<sup>1)</sup>*Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States*

<sup>2)</sup>*Department of Chemical and Petroleum Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, United States*

<sup>3)</sup>*Faculty of Physics, University of Vienna, Kolingasse 14-16, 1090 Vienna, Austria*

(Dated: 22 November 2021)

**Abstract:** Due to the sheer size of chemical and materials space, high throughput computational screening thereof will require the development of new computational methods that are accurate, efficient, and transferable. These methods need to be applicable to electron configurations beyond ground states. To this end, we have systematically studied the applicability of quantum alchemy predictions using a Taylor series expansion on quantum mechanics (QM) calculations for single atoms with different electronic structures arising from different net charges and electron spin multiplicities. We first compare QM method accuracy to experimental quantities including first and second ionization energies, electron affinities, and multiplet spin energy gaps for a baseline understanding of QM reference data. We then investigate the intrinsic accuracy of an approach we call “manual” quantum alchemy schemes compared to the same QM reference data, which employ QM calculations where the basis set of a different element is used for an atom as the limit case of quantum alchemy. We then discuss the reliability of quantum alchemy based on Taylor series approximations at different orders of truncation. Overall, we find that the errors from finite basis set treatments in quantum alchemy are significantly reduced when thermodynamic cycles are employed, which points out a route to improve quantum alchemy in explorations of chemical space. This work establishes important technical aspects that impact the accuracy of quantum alchemy predictions using a Taylor series and provides a foundation for further quantum alchemy studies.

---

<sup>a)</sup>Electronic mail: [jakeith@pitt.edu](mailto:jakeith@pitt.edu)

## I. INTRODUCTION

Productive schemes for the computational screening of molecular compounds<sup>1,2</sup> and materials<sup>3-6</sup> has been an area of interest for many years. Reliable computational predictions of molecular and material properties generally require quantum mechanics (QM) calculations, ideally using the most sophisticated and accurate models possible to account for different physical phenomena. However, the computational cost of schemes involving highly accurate methods also usually their use for predictive insights.<sup>7</sup> Schemes based on Kohn-Sham Density Functional Theory (DFT) are often employed as a useful compromise between computational efficiency and physical accuracy. However, comprehensive screening would ideally consider different atomic-scale compositions, configurations, and system sizes, and the resultant number of combinatorial possibilities makes DFT methods too computationally demanding to be practical. For this reason, many have developed alternative approximate models such as tight binding QM methods,<sup>8-10</sup> cluster expansion methods,<sup>11-15</sup> analytic (reactive) forcefields,<sup>16-18</sup> machine learning potentials,<sup>19</sup> with varying degrees of success.

A promising method applicable for high-throughput screening is quantum alchemy using a Taylor expansion.<sup>20-22</sup> Instead of running many self-consistent field calculations for many target systems by brute force, this method approximates the QM energies for these systems using a Taylor series expansion around the energy of a reference system. This is done by assuming that QM energies of systems are continuously relatable across alchemical space, i.e. across the chemical and materials space where nuclear charge is a variable parameter. Various quantum alchemy methods have been evaluated in predicting molecular properties,<sup>20,21,23,24</sup> bulk material structural properties,<sup>25,26</sup> and surface-adsorbate binding energies and activation energies for catalytic applications.<sup>27-32</sup> To date, most of these studies make use of different quantum alchemy schemes involving distinct implementations of QM on various atomic scale systems that make it non-trivial to assess how and why quantum alchemy using a Taylor series expansion is working overall across chemical space. To achieve a fundamental understanding of the general accuracy and transferability of these Taylor series approximations, we report the performance of quantum alchemy using Taylor series expansions in predicting electronic energies of single atoms and compare to experimentally observable quantities such as ionization energies, electron affinities, and spin multiplet gaps.

Earlier work in this direction<sup>33</sup>, evaluating the alchemical derivatives in the context of conceptual DFT, focused on DFT methods. In this work, we study a correlated wavefunction method, namely the more accurate but more expensive CCSD(T) approach.

In this work, we first report reference calculations that use QM energies from a reasonably high-level wavefunction method with a range of basis sets to show the accuracy of different reference QM calculations when predicting thermodynamic energies that can be compared to experimental data<sup>34–36</sup>. We then report the extent that these energies can also be calculated using “manual” quantum alchemy, which involve alchemical pathways from calculations where nuclear charges have been explicitly modified with respect to a reference system. We then evaluate the efficacy of quantum alchemy based on Taylor series expansions truncated at different orders as a means to approximate the underlying “manual” alchemical pathways. To predict atomic properties including ionization energies, electron affinities, and spin multiplet gaps, we employ thermodynamic cycles and quantify the accuracy of quantum alchemy predictions using a Taylor series expansion. Based on this work, we can achieve a wholistic perspective on how well quantum alchemy using a Taylor series expansion is approximating high-level QM data absolutely on individual atoms, as well as when used in schemes with thermodynamic cycles that would be expected to introduce error cancellations.

## II. METHODS

### A. QM calculations

QM calculations were performed with the open-source Gaussian basis set code, PySCF (v1.7.6)<sup>37–39</sup>. Coupled cluster calculations with single, double, and perturbative triples based on restricted and unrestricted (when necessary) Hartree–Fock (HF) references used default convergence criteria (HF and CCSD were used for one- and two-electron systems, respectively, while CCSD(T) was used for systems containing three electrons or more). We modeled all atoms from H to Ar with charges of 0,  $1\pm$ , and  $2\pm$  when the number of electrons in the system permitted it (Figure 1). We also modeled atomic ground states and a higher energy multiplet state to compute spin multiplicity gaps for each neutral case. Note that some atoms exhibited unstable anions where the electronic energy of the anion was higher than that of the less-negatively charged species. We omit these QM calculations in error

analyses with respect to experiment, but we retain these in later discussions in this work since quantum alchemy might still be useful in predicting QM calculations whether or not they are experimentally justified. For all atoms we tested the aug-cc-pVTZ (aVTZ), aug-cc-pVQZ (aVQZ), and aug-cc-pV5Z (aV5Z) basis sets<sup>40–44</sup>. All aV5Z basis sets were taken from the basis set exchange<sup>45–47</sup>. When appropriate, we compared QM calculations to experimental quantities including first and second ionization energies<sup>34</sup>, electron affinities<sup>35</sup>, and multiplicity gaps<sup>36</sup>. All QM calculations used for quantum alchemy predictions that used a Taylor series expansion converged. A few QM calculations at integer changes in nuclear charge (manual quantum alchemy calculations, see below) did not converge, and these are discussed in the SI (see Table S1). These calculations do not influence the data presented here or the conclusions drawn from the data.

## B. Alchemical potential energy surfaces (PESs)

Alchemical PESs were calculated for all cases above by modifying the nuclear charge ( $Z$ ) of the selected reference atom using manual quantum alchemy calculations. For example, starting from a reference system that was a neutral C atom in a  $^3P$  state, a  $B^-$  atom in the same  $^3P$  state can be obtained by decreasing the nuclear charge on the C atom by one ( $\Delta Z = -1$ ). Alternatively, the neutral C atom could be used as a reference for the  $N^+$  atom in the same  $^3P$  state with a  $\Delta Z = +1$  transmutation. All alchemical transmutations used here will retain the same number of electrons as the reference system. Since practical applications of QM calculations normally use incomplete Gaussian basis sets, a basis set designed for a C atom but used on a  $B^-$  atom will be expected to introduce some error. These basis set errors are expected to decrease when more complete basis sets are used,<sup>48</sup> but we report the extent these errors are present to understand a baseline accuracy for alchemical PESs.

## C. Quantum alchemy predictions using a Taylor series expansion

The energy of a target system can be predicted using quantum alchemy as a Taylor series expansion around the reference system with respect to an arbitrary alchemical pathway

defined by  $\lambda$ ,<sup>21</sup> where  $\lambda = 0$  at the reference and  $\lambda = 1$  at any target:

$$\begin{aligned}
E(\lambda = 1) &= \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n}{\partial \lambda^n} \langle \psi_\lambda | \hat{H}(\lambda = 0) | \psi_\lambda \rangle \Big|_{\lambda=0} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n E(\lambda = 0)}{\partial \lambda^n} \Big|_{\lambda=0} \\
&= E(\lambda = 0) + \sum_{n=1}^{\infty} \frac{1}{n!} \frac{\partial^n E(\lambda = 0)}{\partial \lambda^n} \Big|_{\lambda=0}
\end{aligned} \tag{1}$$

This expression relies on the assumption that the Hellmann-Feynman theorem holds for first order derivatives.<sup>49</sup> From QM, the change in energy with respect to an arbitrary alchemical pathway can be described by nuclear, atomic position, and electronic terms using the chain rule:

$$\frac{\partial E(\lambda)}{\partial \lambda} = \sum_I \frac{\partial E}{\partial N_I} \frac{\partial N_I}{\partial \lambda} + \sum_I \frac{\partial E}{\partial R_I} \frac{\partial R_I}{\partial \lambda} + \frac{\partial E}{\partial N_e} \frac{\partial N_e}{\partial \lambda} \tag{2}$$

where  $\partial N_I$  is the partial differential for nuclear charge of atom  $I$ ,  $\partial R_I$  is the partial differential for position of atom  $I$ , and  $\partial N_e$  is the partial differential for number of electrons.<sup>27</sup> Since this work is only focused on studies of individual atoms that can always be centered at the origin of the coordinate axes, the second term of Equation 2 is zero. As mentioned above, since all alchemical pathways involve a constant number of electrons, the third term of Equation 2 is also zero, which eliminates the dependence of discontinuous derivatives. Thus, our treatment of quantum alchemy only involves energy terms related to so-called *alchemical derivatives* for the change in energy with respect to the change in nuclear charge.

An ongoing concern about using Taylor series expansions is that it may not always converge to a physically meaningful result. However, it has been recently shown that the Taylor series will converge in Hartree-Fock calculations of small molecules.<sup>22</sup> While other mathematically equivalent protocols exist to calculate the energy derivatives from the Taylor series expansion,<sup>32,48</sup> in this work we employ central finite differences<sup>22</sup> with the general formula:

$$\frac{\partial^n E(\lambda)}{\partial \lambda^n} = \sum_{i=0}^n \frac{(-1)^i}{h^n} \binom{n}{i} E\left(\lambda + \left(\frac{n}{2} - i\right)h\right) \tag{3}$$

where  $h$  represents a small displacement from  $\lambda = 0$ . We used  $h = 0.01$  with the findiff<sup>50</sup> Python package and a general accuracy of 2. This requires evaluating the energies for systems at non-integer  $\lambda$  values; in other words, PySCF calculations on a reference atom with a conventional nuclear charge of  $Z$  used nuclear charges of  $Z$ ,  $Z \pm 0.01$ , and  $Z \pm 0.02$ .

Using finite differences, various orders of alchemical derivatives can be used for expressing the Taylor series. The Taylor series at zeroth order would assume that the energy of the reference system itself is a valid approximation for the target system, and so we exclude this order from discussions. By virtue of truncation after the first order being a linear function of the first alchemical derivative, it approximates the alchemical PES as a straight line. Truncations of the Taylor series after the second order in turn approximates the alchemical PES as a second order polynomial, truncations after the third order as a third order polynomial, and truncations after the fourth order as a fourth order polynomial, etc. The use of the finite differences method requires several QM calculations along each alchemical pathway, but it is expected that in computational screening of systems having many possible alchemical transmutation sites, quantum alchemy predictions using a Taylor series will require far fewer calculations overall than that of brute-force investigations.<sup>48</sup>

#### D. Thermodynamic cycles to predict atomic properties

Computational applications of quantum chemistry usually make use of calculation schemes that introduce systematic error cancellations to improve predictive power. Amongst other uses, thermodynamic cycles are a means to introduce error cancellations. Here, thermodynamic cycles were used to calculate first and second ionization energies, electron affinities, and multiplicity gaps for atoms. An example thermodynamic cycle is shown in Figure S3, which can be used to calculate any of these four atomic properties. The top leg (state A to B) represents the reference property calculated using QM calculations ( $\Delta E_1$ ), and the bottom leg (state C to D) represents the target property ( $\Delta E_4$ ) predicted using alchemical transmutation energies for the vertical components:

$$0 = \Delta E_4 - \Delta E_1 - \Delta E_3 + \Delta E_2 \tag{4}$$

$$\Delta E_4 = \Delta E_1 + \Delta E_3 - \Delta E_2 \tag{5}$$

Since vertical legs represent alchemical transmutations, the number of electrons in state A and C must be equal, and the number of electrons in state B and D must be equal. The alchemical transmutation energies  $\Delta E_3$  and  $\Delta E_2$  can be obtained using manual quantum alchemy calculations (obtained with self-consistent field calculations using different nuclear

charges) or quantum alchemy using a Taylor series expansion. Considering the chemical space of this work, several possible reference states could be used for each desired target state. For example, if the aim was to use alchemical treatments to predict the first ionization energy of N ( $N \rightarrow N^+ + e^-$ ,  $\Delta E_1$ ), suitable reference species for  $\Delta E_1$  include:  $C^- \rightarrow C + e^-$ ,  $O^+ \rightarrow O^{2+} + e^-$ , and  $B^{2-} \rightarrow B^- + e^-$ , using transmutations of  $\Delta Z = 1$ ,  $-1$ , and  $2$ , respectively.

### III. RESULTS AND DISCUSSION

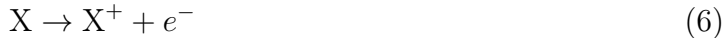
#### A. Comparing QM to experiment

To first assess the accuracy of QM reference data, we used conventional QM calculations to calculate the first and second ionization energies, electron affinities, and excitation energies for atoms H – Ar (when applicable, Figure 1).

Charge		$\leftarrow \Delta Z \rightarrow$																	
	2	H	He	Li	Be	B	C	N	O	F	Ne	Na	Mg	Al	Si	P	S	Cl	Ar
$\uparrow$ 2 <sup>nd</sup> IE	1	H	He	Li	Be	B	C	N	O	F	Ne	Na	Mg	Al	Si	P	S	Cl	Ar
$\uparrow$ 1 <sup>st</sup> IE	0	H	He	Li	Be	B	C	N	O	F	Ne	Na	Mg	Al	Si	P	S	Cl	Ar
$\downarrow$ EA	1-	H	He	Li	Be	B	C	N	O	F	Ne	Na	Mg	Al	Si	P	S	Cl	Ar
	2-	H	He	Li	Be	B	C	N	O	F	Ne	Na	Mg	Al	Si	P	S	Cl	Ar
Reference										Target									

FIG. 1. Illustration of the chemical space sampled in this work. Atoms from H to Ar were studied with different charge states, as labeled on the left. Examples of the first and second ionization energies (IE) and electron affinity (EA) are shown in the solid boxes (target states). Within the thermodynamic cycle, the possible references for each of those targets are also highlighted (dotted boxes). The quantum alchemy schemes we use assume that the number of electrons remains constant between the reference and targets, which limit the possible reference and target combinations.

The first ionization energy describes the energy required to remove an electron from a neutral atom to form a cation:





The second ionization energy describes the energy required to remove an electron from a cation to form a dication:



The first and second ionization energies were calculated for atoms with their ground state multiplicities and compared to experimental values<sup>34</sup>. Table I reports the mean absolute error (MAE), root mean square error (RMSE) and maximum error (max error) for all possible first and second ionization energies calculated using different basis sets (see Figure S1 illustrating errors on an individual-atom basis). As expected, calculated errors decrease as basis set size increases from aVTZ to aV5Z, indicating that more complete basis sets are needed to achieve highest accuracy (Figure S2). Also as expected, errors were generally larger for the second ionization energies than the first ionization energies because the more highly charged dication states are expected to be more challenging to accurately calculate without a tighter core basis set. The maximum errors using the aV5Z basis set were 0.062 eV and 0.540 eV for the first and second ionization energies, respectively. However, errors for aV5Z second ionization energy calculations for individual atoms are all below 0.078 eV except for cases involving Na, Li, and Cl atoms, which have errors  $\geq 0.1$  eV (Figure S1). For achieving higher accuracies, one might consider using even larger basis sets suitable for additional core-valence effects<sup>51,52</sup> or a fully saturated basis set that effectively reaches the complete basis set limit.

TABLE I. Summary of MAE, RMSE, and maximum error by basis set in QM-calculated first and second ionization energies compared to experiment.

IE	Basis Set	MAE (eV)	RMSE (eV)	Max Error (eV)
1st	aVTZ	0.094	0.112	0.210
	aVQZ	0.045	0.056	0.123
	aV5Z	0.024	0.028	0.062
2nd	aVTZ	0.206	0.272	0.794
	aVQZ	0.136	0.215	0.708
	aV5Z	0.103	0.187	0.540

The electron affinity is described as the change in energy to add an electron to a neutral

system:



Electron affinities were calculated and compared to experiment.<sup>35</sup> Some resulting anions (i.e. He, Be, N, Ne, Mg, Ar) do not have experimental records due to their instability,<sup>35</sup> so these data are excluded from the error analysis. As expected, errors in electron affinities decrease as the basis set size increases from aVTZ to aV5Z (Table II and Figure S2). The maximum errors in electron affinities were lower than those obtained from ionization energies, indicating that the augmented correlation consistent basis sets are well suited for calculating electron affinities of atoms (Figure S1).

Basis set	MAE (eV)	RMSE (eV)	Max Error (eV)
aVTZ	0.059	0.078	0.133
aVQZ	0.022	0.029	0.061
aV5Z	0.015	0.019	0.041

TABLE II. Summary of MAE, RMSE, and maximum error by basis set in QM-calculated electron affinities compared to experiment.

Multiplicity gaps were calculated to assess the relative energy difference between a neutral atom’s ground state and an excited state with a different multiplicity (S):

$$X_{ground}(S = y) \rightarrow X_{excited}(S = y \pm 2) \quad (9)$$

For example, the ground state multiplicity of carbon is 3, and the most easily distinguishable excited state with a different electron spin has a multiplicity of 1. Multiplicity gaps were calculated using PySCF by running calculations on ground states and other states having different spin multiplicities, and these values were compared to experiment.<sup>36</sup> Of the available experimental data on excitation energies, we compared this data to experimental excitation data that corresponded best to the output orbital populations from the PySCF calculations (Table S3). Compared to ionization energy and electron affinity data presented thus far, we observe that multiplicity gaps bring the largest errors, but again, errors were lowest overall for the aV5Z (Table III and Figures S1 and S2). Improving the accuracy of calculations of multiplicity gaps would likely require some form of multireference method in lieu of QM calculations using wavefunctions based on a single determinant as well as even more extensive basis sets than those used here.

Basis set	MAE (eV)	RMSE (eV)	Max Error (eV)
aVTZ	0.327	0.515	1.369
aVQZ	0.240	0.328	0.756
aV5Z	0.162	0.220	0.565

TABLE III. Summary of MAE, RMSE, and maximum error by basis set in QM-calculated spin multiplet gaps compared to experiment.

To summarize, we determined errors in atomic energy calculations using CCSD(T) and different basis sets with respect to experimental data. Across all data, the largest basis set, aV5Z, resulted in the most accurate results, and only data using this basis set will be used for the rest of this study. The data reported above should not be considered a paragon of QM calculation data, but it instead demonstrates a useful baseline understanding of errors that would be expected for brute force QM calculations of these atomic properties. We will now report how well related data can be obtained using quantum alchemy predictions that provide a means of circumventing numerous brute force calculations.

## B. Manual quantum alchemy

### 1. *Alchemical PESs*

The alchemical PES relates system energies to their nuclear charges, and this can be reflected as an (arbitrary) alchemical pathway along a variable  $\lambda$ . Two methods to calculate these surfaces are to explicitly do so with QM calculations using variable nuclear charges, which we refer to as “manual” quantum alchemy, or by quantum alchemy using a Taylor series expansion, calculated in this work using finite differences. We first discuss the former.

We calculated the alchemical PESs for all transmutations involving atomic systems with up to 18 electrons considering systems with charges of 0,  $1\pm$ , and  $2\pm$ . For example, the 16 electron cases involved  $\text{Ar}^{2+}$ ,  $\text{Cl}^+$ , neutral S,  $\text{P}^-$ , and  $\text{Si}^{2-}$ . Alchemical PESs relating these systems were constructed by starting from one of these systems (always with the  $^3P$  spin state) and then running new QM calculations with the same basis set but with different nuclear charges assigned (i.e., manual quantum alchemy). This procedure leads to five different alchemical PESs, each using a different element as a reference and that

element’s basis set for all other target species. Figure 2 shows the five alchemical PESs constructed from each reference state, and errors between each target and reference are reported in Figure 3. The same procedure explained above was done for all systems up to and including 18 electrons. For approximately 95% of the cases we studies, alchemical errors increase as  $|\Delta Z|$  increases. (The remaining 5% of the data will be re-investigated in future work.) This outcome is expected because the validity of using a reference system’s basis set for any target system will decrease the larger the difference in nuclear charge from that reference. We also find that alchemical errors are smaller for  $\Delta Z = -1$  than  $\Delta Z = +1$ . Previous work using conceptual DFT in predicting atomic properties had also seen a dependence on transmutation direction.<sup>33</sup> This observed asymmetry in errors was also noted in work by others on diatomic molecules, which showed that the description of different target systems using the basis set of one reference system relied on the direction of  $\Delta Z$ .<sup>48</sup> In the case for atoms, this appears to be due to the augmented finite basis sets for one element generally being slightly more suitable when describing anions than cations of another element. We note that the range of actual errors in an alchemical PES is rather large, ranging from effectively zero to as large as hundreds of eV. Plots depicting these alchemical errors for other reference and target combinations can be generated using our Jupyter Notebooks available free of charge ([github.com/keithgroup/qa-atoms-dimers](https://github.com/keithgroup/qa-atoms-dimers)). We also investigated these errors for Hartree–Fock and correlation energies. Overall, we observe similar ratios of energy to error, suggesting that quantum alchemy impacts both of these energy predictions.

## 2. *Thermodynamic cycles*

As shown above, the errors resulting from manual quantum alchemy calculations can be quite large. Since atomic properties will be calculated using thermodynamic cycles, we next report the performance of manual quantum alchemy calculations when used in thermodynamic cycles (see Figure S3). More specifically, an energy explicitly calculated from QM in the form of an ionization energy, electron affinity, or multiplicity gap (e.g.,  $\Delta E_1$ :  $C \rightarrow C^+ + e^-$ ) will serve as a reference energy. Manual quantum alchemy calculations will then be used to predict other energy contributions (e.g.,  $\Delta E_2$ :  $C \rightarrow N^+$  and  $\Delta E_3$ :  $C^+ \rightarrow N^{2+}$ ) that allow the prediction of a target energy (e.g.,  $\Delta E_4$ :  $N^+ \rightarrow N^{2+} + e^-$ ).

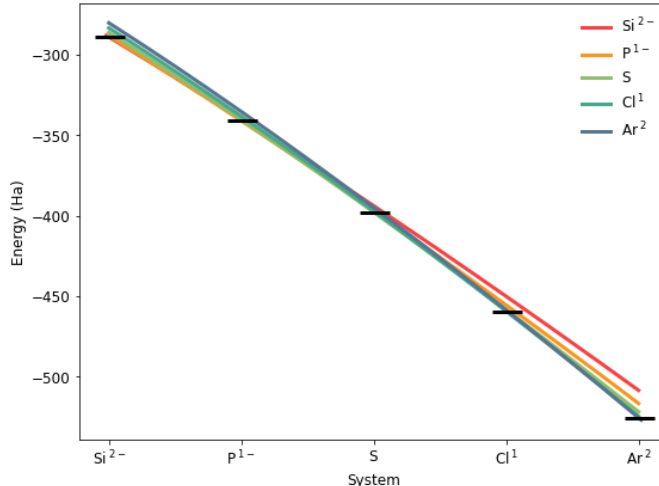


FIG. 2. Alchemical PESs for 16 electron systems as generated from manual quantum alchemy calculations. Each curve makes use of a different reference state and uses self-consistent QM calculations to evaluate energies of the other systems by changing the nuclear charges of the reference to the appropriate target systems. Reference systems are labeled in the legend, and target systems are labeled on the x-axis. Black lines represent the energy of the target system calculated using QM and their own basis set.

Alchemy-based PESs can be used to predict  $\Delta E_2$  and  $\Delta E_3$ , and together  $\Delta E_1$ ,  $\Delta E_2$ , and  $\Delta E_3$  are used to calculate a desired target property  $\Delta E_4$  (see Equation 5). We hypothesized that systematic deficiencies in these predictions would result in error cancellation within the thermodynamic cycle. We now quantify errors in terms of MAE, RMSE, and max error for the first and second ionization energies, electron affinities, and multiplicity gaps using alchemical PESs generated from manual quantum alchemy calculations using thermodynamic cycles (Table IV). Overall, the resulting errors in atomic property predictions from thermodynamic cycles are significantly smaller (MAE  $< 0.6$  eV and max error  $< 5.4$  eV) relative to the alchemical errors discussed above (which were sometimes observed to be two orders of magnitude larger). Specifically, the MAE for thermodynamic cycles is lowest for predictions of the first and second ionization energies (0.143 eV and 0.356 eV, respectively), somewhat higher for electron affinities (0.411 eV), and even higher for multiplicity gaps (0.582 eV). These observations illustrate two main points. First, while PESs calculated using manual quantum alchemy provide fairly significant errors across different systems as shown in Figure 3, when used within thermodynamic cycle schemes, these errors cancel to far less concerning

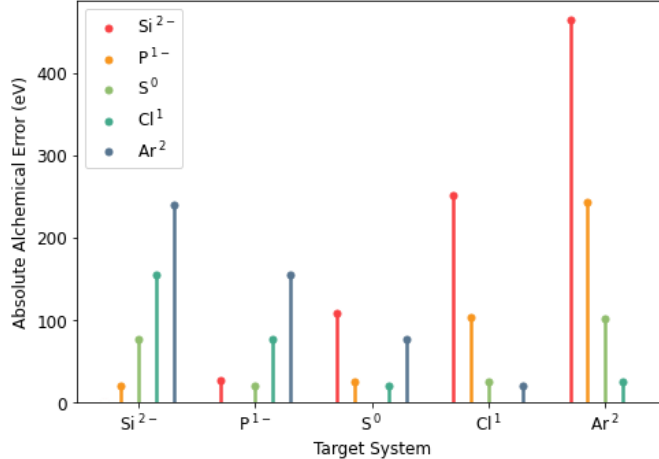


FIG. 3. Plot quantifying the errors in alchemical PESs calculated using manual quantum alchemy for the 16 electron systems shown in Figure 2. Reference systems are labeled in the legend, and target systems are labeled on the x-axis. Alchemical errors are calculated by subtracting the target system energy calculated using the reference’s basis set (i.e., manual quantum alchemy) from the energy of that same target system that was calculated using the correct basis set (i.e., regular QM calculation).

values. Second, the magnitude of these errors are clearly different depending on the charges of the states used in the calculation, and this provides insight into how to improve or at least systematically correct quantum alchemy-based approximations. Again, we see that MAEs increase as  $|\Delta Z|$  increases (Table IV). This is expected since states involving larger changes in nuclear charge with respect to the QM reference calculation are expected to be less accurate, since the basis set used in the quantum alchemical calculation is designed for the reference, not the target. As mentioned above with trends for alchemical PESs, we also see that accuracy depends on the direction of  $\Delta Z$ , as has been observed previously<sup>33</sup> using other quantum alchemy schemes. In most cases, using a reference in which the nuclear charge is reduced to achieve the target system (i.e.,  $\Delta Z = -1$ ) provides the most accurate predictions (again, likely caused by finite basis sets better suited for modeling anions of adjacent atoms rather than cations). This provides insight into what types of references and target combinations may be the most reliably predicted by manual quantum alchemy calculations. For ionization energies, manual quantum alchemy numbers are comparable to DFT accuracy, as shown in Table IV, while the MAE for electron affinities is roughly twice as large compared

to DFT. This points towards electron affinities being more strongly affected by the choice of basis set than ionization energies. In the context of quantum alchemy, this means that more complete basis sets have to be chosen for electron affinities.

Property	$\Delta Z$	MAE (eV)	RMSE (eV)	Max Error (eV)
1st IE	-1	0.038	0.074	0.219
	1	0.074	0.179	0.716
	2	0.320	0.756	2.929
	Overall	0.143	0.447	2.929
	PBE/def2-QZVP	0.174	0.224	0.441
2nd IE	1	0.126	0.197	0.498
	2	0.210	0.290	0.787
	3	0.758	1.341	4.589
	Overall	0.356	0.785	4.589
EA	-2	0.459	1.092	3.814
	-1	0.370	1.064	3.861
	1	0.406	0.917	3.135
	Overall	0.411	1.053	3.861
	PBE/def2-QZVP	0.189	0.217	0.335
MG	-2	0.558	1.222	4.286
	-1	0.564	1.146	3.551
	1	0.495	0.869	2.015
	2	0.715	1.556	5.312
	Overall	0.582	1.218	5.312

TABLE IV. Summary of MAE, RMSE, and maximum error for first and second ionization energies, electron affinities, and multiplicity gaps calculated using manual quantum alchemy with respect to QM. DFT comparison values for PBE/def2-QZVP are shown for the subset of systems included in the GMTKN55 data set<sup>53</sup>.

## C. Quantum alchemy using a Taylor series expansion

### 1. Alchemical PESs

Quantum alchemy using a Taylor series expansion can also be used to approximate alchemical PESs. Unlike manual quantum alchemy, which uses explicit QM calculations using the basis set of the reference to construct the PES, quantum alchemy using a Taylor series simply relies on a few QM calculations at and near the reference state to estimate the PES using finite differences. To continue with the example used above, alchemical PESs can be approximated using a Taylor series for each 16 electron system as a reference. Figure 4 shows Taylor series-approximated PESs when  $\text{Si}^{2-}$ , S, or  $\text{Ar}^{2+}$  are used as the reference. These plots also depict the performance of various truncations of the Taylor series. This example depicts the ability of higher orders, specifically second and third orders, to accurately predict alchemical PESs. While it may appear that the most accurate Taylor series-predicted PES uses S as a reference, note that this PES was constructed for  $\Delta Z$  values up to and including  $\pm 2$ , whereas the PESs predicted using  $\text{Si}^{2-}$  or  $\text{Ar}^{2+}$  as references contain  $|\Delta Z|$  up to and including 4.

We find that the Taylor series at the fourth order deviates from the quantum alchemical PES near  $\Delta Z \geq 2$  when  $\text{Si}^{2-}$  is used as a reference. Interestingly, this order does not significantly deviate from the PES, even at larger  $\Delta Z$  values, when  $\text{Ar}^{2+}$  is used as a reference. These observations highlight two possible sources of error in predictions of PESs using a Taylor series. First, these observations support our above discussions that the sign of  $\Delta Z$  can significantly impact predictions. As observed in Figure 2, more deviation in manual quantum alchemy-calculated PESs is observed for the right side of the plot (for positive  $\Delta Z$ s) than the left side of the plot (for negative  $\Delta Z$ s). When the nuclear charge of the reference is decreased (negative  $\Delta Z$ ), quantum alchemy predictions using a Taylor series are more accurate, in line with earlier observations for mean-field calculations<sup>33</sup>. Second, the deviations in Taylor series-predicted PESs are only observed here in for the fourth order. This indicates that while the sign of  $\Delta Z$  is important in the accuracy of quantum alchemy using a Taylor series, another factor is at play. Lower orders of the Taylor series accurately capture the alchemical PESs using  $\text{Si}^{2-}$  as a reference, which suggests that observed errors in the fourth order predicted PESs may be attributed to insufficient numerical precision



and/or insufficiently tight convergence of reference system calculations that are required for higher order derivatives in the Taylor series to be reliable.<sup>22</sup> Other approaches to evaluate higher order alchemical derivatives such as the conceptual DFT route face similar numerical issues from the fourth order onwards<sup>33</sup>.

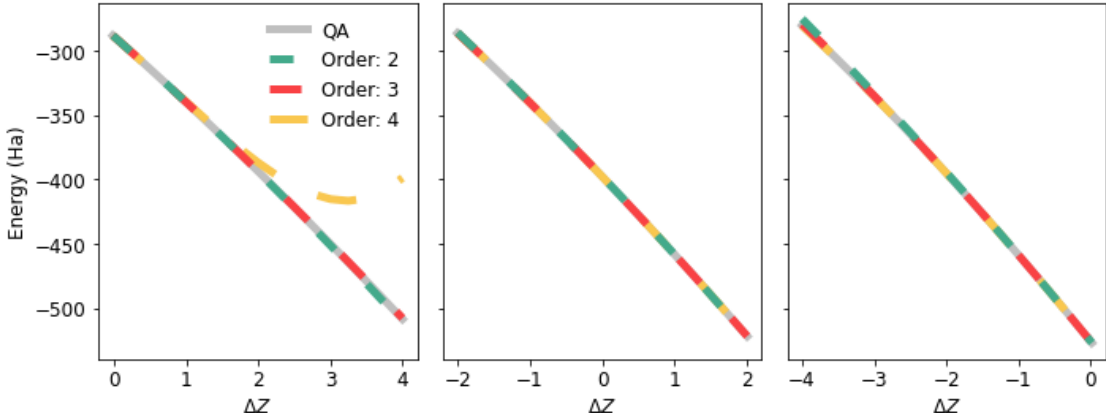


FIG. 4. Comparison of predicted PESs using a Taylor series expansion truncated at various orders for (left)  $\text{Si}^{2-}$ , (middle) S, or (right)  $\text{Ar}^{2+}$  as a reference. Gray solid line indicates alchemical PES calculated using manual quantum alchemy (QA).

Recall that PESs calculated using manual quantum alchemy resulted in large errors with respect to QM-calculated PESs. These same errors will manifest in the Taylor series-predicted PESs when compared to QM calculations. However, quantum alchemy using a Taylor series expansion predicts the alchemical PES calculated using manual quantum alchemy. In other words, the alchemical PESs predicted by the Taylor series rely on the basis set of the reference, similar to how the PES can be constructed using manual quantum alchemy calculations where the nuclear charge of the reference was altered. Therefore, Taylor series-predicted PESs in Figure 4 are compared to manual quantum alchemy-calculated PESs. Qualitatively, we find good agreement between higher orders of the Taylor series-predicted PESs and manual quantum alchemy-calculated PESs. Quantitatively, second and third order truncations of the Taylor series provide the most accurate predictions of the PESs calculated using manual quantum alchemy (Figure 5). These results provide insight into how the Taylor series expansion calculated using finite differences may impose its own error when predicting alchemical PESs, where as expected, increased errors are observed as the magnitude of  $|\Delta Z|$  increases. However, our main goal is to assess any errors in Tay-

lor series-predicted atomic properties when using thermodynamic cycles, which is discussed next.

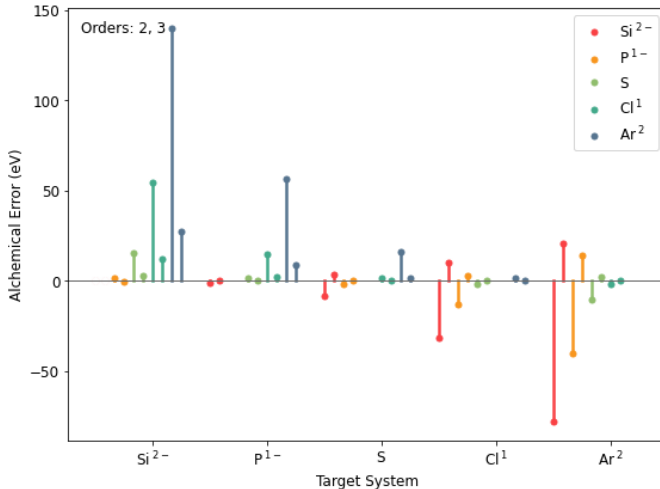


FIG. 5. Plot quantifying the errors in PESs predicted by second and third order Taylor series expansions for the 16 electron systems shown in Figure 2. Reference systems are labeled in the legend, and target systems are labeled on the x-axis. Errors are with respect to PESs calculated using manual quantum alchemy.

## 2. Thermodynamic Cycles

To assess the performance of quantum alchemy predictions of atomic properties using a Taylor series expansion, we now report errors in these Taylor series predictions at various truncations of the expansion with respect to QM-calculated values. All predictions here use thermodynamic cycles to predict atomic properties (see Figure S3). Within this framework,  $\Delta E_2$  and  $\Delta E_3$  are estimated using a Taylor series expansion about the references A and B, respectively. For first and second ionization energy and electron affinity predictions, we find that second order predictions provide the lowest overall MAE (Table V). There are two reasons why the second order Taylor series approximation likely leads to the most accurate results overall. First, the alchemical PESs resemble a relatively shallow parabola (Figure 2), which is modeled sufficiently well as a second-order polynomial rather than third- and fourth-order polynomials predicted from finite differences calculations. Second, higher order predictions were recently shown to require high numerical precision and tight convergence

criteria,<sup>22</sup> suggesting that in this work, higher orders require derivatives to be calculated with much higher precision. However, since errors for the second order predictions are relatively low, the current numerical precision used to calculate Taylor series derivatives may be acceptable up to this order. Further, if greater numerical precision is used, this study suggests that we may be able to further decrease errors associated with quantum alchemy using a Taylor series and approach QM accuracies. The first order Taylor series approximation provides the most accurate predictions compared to QM-calculated properties for multiplicity gaps (Table V). However, errors significantly increase as the Taylor series order increases. We suspect that a large fraction of these errors are due to employing a Taylor series expansion of QM calculations that did not employ multireference methods, not due to the Taylor series expansion itself. Further, errors in the Taylor series are uncontrolled, meaning that numerical noise in the prediction of one leg of the thermodynamic cycle may not necessarily occur in the other vertical leg. Thus, some errors at higher order predictions using the Taylor series may be caused by uncontrolled errors. Overall, errors due to quantum alchemy predictions using a Taylor series expansion are higher than that of errors calculated using manual quantum alchemy when compared to QM, which suggests that the Taylor series expansion introduces some error of its own.

We cannot compare overall errors for each atomic property predicted by quantum alchemy using a Taylor series because each atomic property has a unique set of  $\Delta Z$  values that were explored within the scope of this study. This is most easily depicted in Figure 1. For example, first ionization energies used references whose nuclear charges were changed by either  $\Delta Z = -1, 1, \text{ or } 2$ , whereas second ionization energies used references whose nuclear charges were changed by either  $\Delta Z = 1, 2, \text{ or } 3$ . Electron affinities used references whose nuclear charges were changed by either  $\Delta Z = -2, -1, \text{ or } 1$ , and multiplicity gaps used references whose nuclear charges were changed by either  $\Delta Z = -2, -1, 1, \text{ or } 2$ . Therefore, we can better understand the performance and errors of the Taylor series approximation by considering the errors as a function of  $\Delta Z$  (Tables S4, S5, and S6 report errors by  $\Delta Z$  for all Taylor series orders and Table VI reports errors by  $\Delta Z$  for the most accurate Taylor series order). When the errors are broken down by  $\Delta Z$ , depending on the atomic property, the most accurate predictions appear to depend on the order of the Taylor series approximation (see Tables S4, S5, and S6). To be consistent across atomic properties and  $\Delta Z$ s, we will only compare errors in the Taylor series approximation using the second order for ionization energies and electron

Property	Taylor Series Order	MAE (eV)	RMSE (eV)	Max Error (eV)
1st IE	0	14.917	16.576	50.521
	1	5.328	6.728	16.978
	2	2.948	6.045	21.682
	3	10.636	42.447	284.473
	4	181.430	726.470	4.928x10 <sup>3</sup>
2nd IE	0	26.061	28.941	74.353
	1	12.889	17.025	54.106
	2	7.756	14.621	52.571
	3	39.957	149.165	980.702
	4	989.062	3.823x10 <sup>3</sup>	2.541x10 <sup>4</sup>
EA	0	15.155	20.144	74.353
	1	6.222	10.411	54.190
	2	1.116	1.629	3.835
	3	1.641	5.317	34.304
	4	11.464	43.040	291.468
MG	0	12.451	26.520	137.316
	1	6.579	14.677	74.357
	2	447.857	2.189x10 <sup>3</sup>	1.643x10 <sup>4</sup>
	3	2.289x10 <sup>4</sup>	9.948x10 <sup>4</sup>	5.650x10 <sup>5</sup>
	4	4.341x10 <sup>6</sup>	2.189x10 <sup>7</sup>	1.644x10 <sup>8</sup>

TABLE V. Summary of MAE, RMSE, and maximum error for first and second ionization energies, electron affinities, and multiplicity gaps as predicted using a Taylor series expansion truncated at various orders with respect to QM and using thermodynamic cycles.

affinities and using the first order for multiplicity gaps.

Based on the second order Taylor series errors for ionization energies and electron affinities shown in Table VI as a function of  $\Delta Z$ , we have identified three main sources of error

Property <sup>a</sup>	$\Delta Z$	MAE (eV)	RMSE (eV)	Max Error (eV)
1st IE	-1	0.085	0.149	0.480
	1	0.688	0.902	2.745
	2	8.213	10.537	21.682
2nd IE	1	0.152	0.222	0.576
	2	3.570	3.891	6.512
	3	20.332	25.566	52.571
EA	-2	0.911	1.329	3.567
	-1	0.456	0.977	3.393
	1	2.023	2.311	3.835
MG	-2	12.765	25.183	74.357
	-1	3.413	6.517	19.263
	1	2.892	5.112	12.519
	2	7.864	14.031	45.548

TABLE VI. Summary of MAE, RMSE, and maximum error for first and second ionization energies, electron affinities, and multiplicity gaps as predicted using a Taylor series expansion and thermodynamic cycles with respect to QM as a function of  $\Delta Z$ .

<sup>a</sup>Reported Taylor series data for second order (IE and EA) and first order (MG) truncations.

in predictions when using a Taylor series expansion for quantum alchemy: 1) magnitude of change in nuclear charge from reference to target (magnitude of  $|\Delta Z|$ ), 2) direction of change in nuclear charge from the reference to target (sign of  $\Delta Z$ ), and 3) intrinsic uncertainties associated with reference systems. Some of these trends do not hold for Taylor series predictions of multiplicity gaps, which may be due to the first order truncation or a result of the overall large errors. As mentioned previously, we attribute these large errors to using single reference QM methods for the Taylor series expansion and since higher numerical precision is likely required for higher orders of the Taylor series. Even with the current numerical settings, the results in Table VI are comparable to DFT accuracy as reported in Table IV

for  $\Delta Z$  of -1.

Generally, Taylor series errors increase as the magnitude of  $|\Delta Z|$  increases. Considering the Taylor series expansion, this observation is expected and has been noted in previous studies.<sup>48</sup> When the Taylor series is truncated after the second order, the performance will be limited more locally to the reference system. As higher order corrections are calculated reasonably, this error is expected to decrease, as has been shown recently.<sup>22</sup>

Our results also suggest that more work is needed to establish if the sign of  $\Delta Z$  truly impacts the intrinsic accuracy of predictions in quantum alchemy using a Taylor series expansion. As discussed above, negative  $\Delta Z$ s usually provide more accurate predictions than their positive counterparts, but this may be due to systematically smaller errors in anions with the finite basis sets used here. Since our work focuses on single atoms, we hypothesize that errors that are dependent on  $\Delta Z$  sign likely arise because incomplete basis sets are used, which are not designed for handling nuclear charge changes or optimized for charged systems. This trend was also observed in the manual quantum alchemy data, further supporting this hypothesis (Figure S4). These errors are expected to decrease significantly (and thus the difference in errors from positive and negative  $\Delta Z$ s are expected to decrease) as one approaches the complete basis set limit.

In addition to errors identified pertaining to the magnitude of  $|\Delta Z|$  and the direction of  $\Delta Z$ , we also find evidence of a third source of error in predictions using a Taylor series, which involves the intrinsic uncertainty in QM calculations of the reference. Specifically, increased errors are observed when dianionic references are used. For example, second order Taylor series predictions of electron affinities indicate that the MAE for  $\Delta Z = -2$  is smaller than for  $\Delta Z = 1$ . This observation is unexpected, as smaller  $|\Delta Z|$  values usually provide more accurate predictions than larger ones. We attribute this observation to the use of a dianionic reference whose nuclear charge is altered by 1. While dianions can be calculated using QM, none of these reported here have been observed experimentally. Our results indicate that unstable references should not be used for quantum alchemy predictions, as it is difficult to extrapolate from an unstable reference.

## IV. CONCLUSIONS

We have quantified errors arising from manual quantum alchemy calculations and quantum alchemy predictions using a Taylor series expansion for atomic alchemical PESs and properties. We find that significant errors can arise when a manual quantum alchemy calculations are used, which use a reference atom’s basis set and then alter the nuclear charge to represent a target atom, some of which can be hundreds of eV. However, we show that when atomic properties are calculated using manual quantum alchemy and thermodynamic cycles, these errors are significantly reduced ( $\text{MAE} < 0.6 \text{ eV}$ ). This same concept is applied to quantum alchemy predictions of atomic properties using a Taylor series expansion, where error cancellation is observed, leading to more accurate results. While this thermodynamic scheme has been used previously for other quantum alchemy predictions, our atomic studies have provided a clear understanding as to why predictions using a Taylor series have been quite promising in thermodynamic cycles. For ionization energies and electron affinities, we find that the second order Taylor series predictions are the most accurate, and for multiplicity gaps, we find that the first order Taylor series predictions are the most accurate. We attribute these findings to the fact that, while higher orders of the Taylor series expansion should be more accurate, higher numerical precision and tighter convergence is required to observe this in practice. Further, we identify that for quantum alchemy using a Taylor series, there are three significant sources of error: the magnitude of the change in nuclear charge from the reference to the target, the sign of the change in nuclear charge from the reference to the target, and the intrinsic uncertainty of the QM-calculated reference. These sources of error easily explain errors in ionization energies and electron affinities, but these errors do not robustly explain trends in first order Taylor series predictions of multiplicity gaps. However, we note that errors can be quite large for multiplicity gap predictions and that first order Taylor series predictions provide the most accurate results, which may explain why these errors deviate from the sources of error and trends discussed here. Taylor series predictions are more accurate for smaller  $|\Delta Z|$  values, which is expected due to the nature of the Taylor series expansion at second order approximations. The sign of the nuclear charge change is important in the accuracies of Taylor series predictions, and we find that more accurate predictions are provided when the nuclear charge of the reference is decreased to the target. This directionality is attributed to the use of finite basis sets, where quantum alchemy pre-

dictions are likely to generally be dependent on direction. Further, this directionality may be leveraged so that reference’s are strategically chosen based on the desired target systems. Lastly, our results suggest that the intrinsic uncertainty of reference calculations can impact predictions using a Taylor series expansion. These errors should decrease as higher order corrections are used, given the necessary numerical precision and tight convergence criteria are reached,<sup>22</sup> providing a positive outlook for using a Taylor series expansion. Taken together, this work highlights fundamental areas that accuracies from predictions obtained via quantum alchemy using a Taylor series expansion depend on and provide a foundation for understanding more complicated systems involving atomic interactions.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available on GitHub at [github.com/keithgroup](https://github.com/keithgroup), repositories `qa-atoms-data` and `qa-atoms-dimers`.

## ACKNOWLEDGMENTS

We thank Professors Ken Jordan and Florian Libisch for helpful discussions over the course of this work. E.A.E. acknowledges support from the Hurd W. Safford Graduate Teaching Fellowship and the Kenneth P. Dietrich School of Arts & Sciences Graduate Fellowship. We acknowledge financial support from the U.S. National Science Foundation: GRFP-1747452 (for C.D.G.) and CHE-1856460 (J.A.K.) and the R. K. Mellon foundation and the University of Pittsburgh’s Swanson School of Engineering. Calculations were run using resources provided by the University of Pittsburgh Center for Research Computing.

## REFERENCES

- <sup>1</sup>D. B. Kitchen, H. Decornez, J. R. Furr, and J. Bajorath, “Docking and scoring in virtual screening for drug discovery: Methods and applications,” *Nat. Rev. Drug Discovery* **3**, 935–949 (2004).
- <sup>2</sup>S. J. Y. Macalino, V. Gosu, S. Hong, and S. Choi, “Role of computer-aided drug design in modern drug discovery,” *Arch. Pharmacol Res.* **38**, 1686–1701 (2015).



- <sup>3</sup>J. K. Nørskov, T. Bligaard, J. Rossmeisl, and C. H. Christensen, “Towards the computational design of solid catalysts,” *Nat. Chem.* **1**, 37–46 (2009).
- <sup>4</sup>Y. J. Colón and R. Q. Snurr, “High-throughput computational screening of metal–organic frameworks,” *Chem. Soc. Rev.* **43**, 5735–5749 (2014).
- <sup>5</sup>A. Jain, Y. Shin, and K. A. Persson, “Computational predictions of energy materials using density functional theory,” *Nat. Rev. Mater.* **1**, 1–13 (2016).
- <sup>6</sup>Q. Wang, L. Velasco, B. Breitung, and V. Presser, “High-entropy energy materials in the age of big data: A critical guide to next-generation synthesis and applications,” *Adv. Energy Mater.* , 2102355 (2021).
- <sup>7</sup>J. A. Keith, V. Vassilev-Galindo, B. Cheng, S. Chmiela, M. Gastegger, K.-R. Müller, and A. Tkatchenko, “Combining machine learning and computational chemistry for predictive insights into chemical systems,” *Chem. Rev.* **121**, 9816–9872 (2021).
- <sup>8</sup>Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim, and M. Karplus, “A qm/mm implementation of the self-consistent charge density functional tight binding (scc-dftb) method,” *J. Phys. Chem. B* **105**, 569–585 (2001).
- <sup>9</sup>B. Aradi, B. Hourahine, and T. Frauenheim, “Dftb+, a sparse matrix-based implementation of the dftb method,” *J. Phys. Chem. A* **111**, 5678–5684 (2007).
- <sup>10</sup>M. Gaus, Q. Cui, and M. Elstner, “Dftb3: Extension of the self-consistent-charge density-functional tight-binding method (scc-dftb),” *J. Chem. Theory Comput.* **7**, 931–948 (2011).
- <sup>11</sup>D. J. Schmidt, W. Chen, C. Wolverton, and W. F. Schneider, “Performance of cluster expansions of coverage-dependent adsorption of atomic oxygen on pt(111),” *J. Chem. Theory Comput.* **8**, 264 (2012).
- <sup>12</sup>T. Rehman, M. Jaipal, and A. Chatterjee, “A cluster expansion model for predicting activation barrier of atomic processes,” *J. Comput. Phys.* **243**, 244 (2013).
- <sup>13</sup>S. Verma, T. Rehman, and A. Chatterjee, “A cluster expansion model for rate constants of surface diffusion processes on ag, al, cu, ni, pd and pt(100) surfaces,” *Surf. Sci.* **613**, 114 (2013).
- <sup>14</sup>L. M. Herder, J. M. Bray, and W. F. Schneider, “Comparison of cluster expansion fitting algorithms for interactions at surfaces,” *Surf. Sci.* **640**, 104 (2015).
- <sup>15</sup>A. Bajpai, K. Frey, and W. F. Schneider, “Binary approach to ternary cluster expansions: No–o–vacancy system on pt(111),” *J. Phys. Chem. C* **121**, 7344 (2017).

- <sup>16</sup>A. P. Sutton and J. Chen, “Long-range finnis–sinclair potentials,” *Philos. Mag. Lett.* **61**, 139–146 (1990).
- <sup>17</sup>J. Yu, S. B. Sinnott, and S. R. Phillpot, “Charge optimized many-body potential for the si/sio<sub>2</sub> system,” *Phys. Rev. B* **75**, 085311 (2007).
- <sup>18</sup>T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, and A. C. T. van Duin, “The reaxff reactive force-field: Development, applications and future directions,” *npj Comput. Mater.* **2**, 1–14 (2016).
- <sup>19</sup>O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, “Machine learning force fields,” *Chem. Rev.* **121**, 10142–10186 (2021).
- <sup>20</sup>O. A. von Lilienfeld, “Accurate ab initio energy gradients in chemical compound space,” *J. Chem. Phys.* **131**, 164102 (2009).
- <sup>21</sup>G. F. von Rudorff and O. A. von Lilienfeld, “Alchemical perturbation density functional theory,” *Phys. Rev. Res.* **2**, 023220 (2020).
- <sup>22</sup>G. F. von Rudorff, “Arbitrarily accurate quantum alchemy,” *J. Chem. Phys.* (2021), 10.1063/5.0073941.
- <sup>23</sup>G. F. von Rudorff and O. A. von Lilienfeld, “Atoms in molecules from alchemical perturbation density functional theory,” *J. Phys. Chem. B* **123**, 10073–10082 (2019).
- <sup>24</sup>G. F. von Rudorff and O. A. von Lilienfeld, “Rapid and accurate molecular deprotonation energies from quantum alchemy,” *Phys. Chem. Chem. Phys.* **22**, 10519–10525 (2020).
- <sup>25</sup>M. to Baben, J. O. Achenbach, and O. A. von Lilienfeld, “Guiding ab initio calculations by alchemical derivatives,” *J. Chem. Phys.* **144**, 104103 (2016).
- <sup>26</sup>A. Solovyeva and O. A. von Lilienfeld, “Alchemical screening of ionic crystals,” *Phys. Chem. Chem. Phys.* **18**, 31078–31091 (2016).
- <sup>27</sup>D. Sheppard, G. Henkelman, and O. A. von Lilienfeld, “Alchemical derivatives of reaction energetics,” *J. Chem. Phys.* **133**, 084104 (2010).
- <sup>28</sup>R. Balawender, M. A. Welearegay, M. Lesiuk, F. De Proft, and P. Geerlings, “Exploring chemical space with the alchemical derivatives,” *J. Chem. Theory Comput.* **9**, 5327–5340 (2013).
- <sup>29</sup>K. Saravanan, J. R. Kitchin, O. A. von Lilienfeld, and J. A. Keith, “Alchemical predictions for computational catalysis: Potential and limitations,” *J. Phys. Chem. Lett.* **8**, 5002–5007

- (2017).
- <sup>30</sup>R. Balawender, M. Lesiuk, F. De Proft, and P. Geerlings, “Exploring chemical space with alchemical derivatives: Bn-simultaneous substitution patterns in c<sub>6</sub>0,” *J. Chem. Theory Comput.* **14**, 1154–1168 (2018).
- <sup>31</sup>C. D. Griego, K. Saravanan, and J. A. Keith, “Benchmarking computational alchemy for carbide, nitride, and oxide catalysts,” *Adv. Theory Simul.* **2**, 1800142 (2019).
- <sup>32</sup>C. D. Griego, J. R. Kitchin, and J. A. Keith, “Acceleration of catalyst discovery with easy, fast, and reproducible computational alchemy,” *Int. J. Quantum Chem.* **121**, e26380 (2020).
- <sup>33</sup>R. Balawender, M. Lesiuk, F. D. Proft, C. V. Alsenoy, and P. Geerlings, “Exploring chemical space with alchemical derivatives: Alchemical transformations of h through ar and their ions as a proof of concept,” *Phys. Chem. Chem. Phys.* **21**, 23865–23879 (2019).
- <sup>34</sup>J. E. Sansonetti and W. C. Martin, “Handbook of basic atomic spectroscopic data,” *J. Phys. Chem. Ref. Data* **34**, 1559–2259 (2005).
- <sup>35</sup>W. M. Haynes, D. R. Lide, and T. J. Bruno, *CRC Handbook of Chemistry and Physics*, edited by W. M. Haynes, Vol. 2016–2017, 97th edition (CRC Press, 2017).
- <sup>36</sup>A. Kramida, Y. Ralchenko, J. Reader, and NIST ASD Team, NIST Atomic Spectra Database (ver. 5.8), [Online]. Available: <https://physics.nist.gov/asd> [2021, October 20]. National Institute of Standards and Technology, Gaithersburg, MD. (2020).
- <sup>37</sup>Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, *et al.*, “Recent developments in the PySCF program package,” *J. Chem. Phys.* **153**, 024109 (2020).
- <sup>38</sup>Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, *et al.*, “Pyscf: the python-based simulations of chemistry framework,” *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **8**, e1340 (2018).
- <sup>39</sup>Q. Sun, “Libcint: An efficient general integral library for gaussian basis functions,” *J. Comp. Chem.* **36**, 1664–1671 (2015).
- <sup>40</sup>B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, and A. K. Wilson, “Gaussian basis sets for use in correlated molecular calculations. vii. valence, core-valence, and scalar relativistic basis sets for li, be, na, and mg,” *Theor. Chem. Acc.* **128**, 69–82 (2011).
- <sup>41</sup>D. E. Woon and T. H. Dunning, “Gaussian basis sets for use in correlated molecular calculations. iv. calculation of static electrical response properties,” *J. Chem. Phys.* **100**,

- 2975–2988 (1994).
- <sup>42</sup>D. E. Woon and T. H. Dunning, “Gaussian basis sets for use in correlated molecular calculations. iii. the atoms aluminum through argon,” *J. Chem. Phys.* **98**, 1358–1371 (1993).
- <sup>43</sup>R. A. Kendall, T. H. Dunning, and R. J. Harrison, “Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions,” *J. Chem. Phys.* **96**, 6796–6806 (1992).
- <sup>44</sup>T. H. Dunning, “Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen,” *J. Chem. Phys.* **90**, 1007–1023 (1989).
- <sup>45</sup>B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibbsom, and T. L. Windus, “A new basis set exchange: An open, up-to-date resource for the molecular sciences community,” *J. Chem. Inf. Model.* **59**, 4814–4820 (2019).
- <sup>46</sup>D. Feller, “The role of databases in support of computational chemistry calculations,” *J. Comput. Chem.* **17**, 1571–1586 (1996).
- <sup>47</sup>K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, “Basis set exchange: A community database for computational sciences,” *J. Chem. Inf. Model.* **47**, 1045–1052 (2007).
- <sup>48</sup>G. Domenichini, G. F. von Rudorff, and O. A. von Lilienfeld, “Effects of perturbation order and basis set on alchemical predictions,” *J. Chem. Phys.* **153**, 144118 (2020).
- <sup>49</sup>R. P. Feynman, “Forces in molecules,” *Phys. Rev.* **56**, 340–343 (1939).
- <sup>50</sup>M. Baer, Maroba/findiff: A Python package for finite difference numerical derivatives and partial differential equations in any number of dimensions.
- <sup>51</sup>D. E. Woon and T. H. Dunning Jr, “Gaussian basis sets for use in correlated molecular calculations. v. core-valence basis sets for boron through neon,” *J. Chem. Phys.* **103**, 4572–4585 (1995).
- <sup>52</sup>K. A. Peterson and T. H. Dunning Jr, “Accurate correlation consistent basis sets for molecular core–valence correlation effects: The second row atoms al–ar, and the first row atoms b–ne revisited,” *J. Chem. Phys.* **117**, 10548–10560 (2002).
- <sup>53</sup>L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi, and S. Grimme, “A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions,” *Phys. Chem. Chem. Phys.* **19**, 32184–32215 (2017).