

# Surface Segregation Studies in Ternary Noble Metal Alloys: Comparing DFT and Machine Learning with Experimental Data

Kirby Broderick,<sup>[a]</sup> Robert A. Burnley,<sup>[a]</sup> Andrew J. Gellman,<sup>[a]</sup> and John R. Kitchin<sup>\*[a]</sup>

Surface segregation, whereby the surface composition of an alloy differs systematically from the bulk, has historically been hard to study, because it requires experimental and modeling methods that span alloy composition space. In this work, we study surface segregation in catalytically relevant noble and platinum-group metal alloys with a focus on three ternary systems: AgAuCu, AuCuPd, and CuPdPt. We develop a data set of 2478 fcc slabs with those compositions including all three low-index crystallographic orientations relaxed with Density Functional Theory using the PBEsol functional with D3 dispersion corrections. We fine-tune a machine learning model on this data and use the model in a series of 1800 Monte Carlo

simulations spanning ternary composition space for each surface orientation and ternary chemical system. The results of these simulations are validated against prior experimental surface segregation data collected using composition spread alloy films for AgAuCu and AuCuPd. Our findings reveal that simulations conducted using the (110) orientation most closely match experimentally observed surface segregation trends, and while predicted trends qualitatively match observation, biases in the PBEsol functional limit numeric accuracy. This study advances understanding of surface segregation and the utility of computational studies and highlights the need for further improvements in simulation accuracy.

## Introduction

Throughout history, alloys such as bronze and steel have played an important role in commerce and industry due to their mechanical and chemical properties. In modern surface science, catalysts may be alloyed to take advantage of novel combinations of electronic or structural properties to maximize activity, selectivity, or stability while minimizing the use of precious metals.<sup>[1]</sup> Because catalysis is an interfacial phenomenon, the structure and composition of the alloy surface are critical in determining whether alloying will have catalytically favorable effects.

The breadth of elements and structures available to make catalytic materials creates a significant combinatorial search space for scientists seeking to optimize catalyst activity. This challenge is only compounded by the requirement that catalytically active surfaces be stable and active under reaction conditions. Meanwhile, although the scope of experiments is increasing as a consequence of the digital revolution, the high-quality experimental data sets and characterization needed to scale to industrial processes remain orders of magnitude smaller

than the feasible search space. To address this disparity, scientists may turn to computational tools that use physics to predict properties of materials and interfaces to augment experimental approaches. In heterogeneous catalysis, Density Functional Theory (DFT) has frequently been found to reasonably balance computational speed and physical accuracy for modeling crystalline material surfaces.<sup>[2]</sup> More recently, computational chemists have looked to machine learning using large data sets as a way to further accelerate discoveries with minimal loss in accuracy.<sup>[3,4]</sup>

However, even as the accuracy of machine learned DFT surrogate models improves, important challenges remain in applying these models to understanding or supplementing real-world data to achieve rational materials design.<sup>[5]</sup> One challenge for accurate materials simulations is the need to model disorder, which is computationally costly and can stymie materials discovery efforts that assume materials are ordered crystals.<sup>[6]</sup> A challenge more specific to ab initio catalyst design is the need to understand surface stability. Surface structure and area are crucial determinants of catalyst activity, and while laboratory experiments can utilize single-crystal materials with known surface area and geometry, real-world catalysts are polymorphic nanoparticles that express many surfaces and surface orientations that may differ dramatically in their ability to catalyze a target reaction.<sup>[7]</sup> The stability of these nanoparticles also depends in part on the stability of their surfaces.<sup>[8]</sup> This poses a dual inference problem for computational representations: models must simultaneously predict structures and their exposed surfaces. While this problem pervades interfacial materials design, it is especially challenging in catalysis because of the need to consider interacting phenomena on both sides of the interface simultaneously.<sup>[9,10]</sup>

[a] K. Broderick, R. A. Burnley, A. J. Gellman, J. R. Kitchin  
Carnegie Mellon University Department of Chemical Engineering,  
5000 Forbes Ave, Pittsburgh, Pennsylvania 15213, United States  
E-mail: jkitchin@andrew.cmu.edu

Supporting information for this article is available on the WWW under  
<https://doi.org/10.1002/cphc.202400073>

© 2024 The Authors. ChemPhysChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

One example of the unique complexity that catalysis poses for surface modeling is adsorbate-induced surface segregation. In multicomponent materials, one element is often preferentially enriched at the surface relative to the bulk due to differing surface energy or lattice mismatch. If the surface has significant chemical interaction with its environment, chemisorbed molecules can exert substantial influence over surface composition. However, the chemical composition of the surface also influences the quantity and composition of adsorbates. This phenomenon has been shown to modify catalyst activity in real materials,<sup>[11,12]</sup> and computational studies have also modeled these effects using Monte Carlo methods with DFT-based potentials.<sup>[13–15]</sup>

### Machine-Learned DFT Surrogates

While DFT can assay materials faster than conventional experiments, its computational cost and poor scaling with electron count imposes practical constraints on simulation size and quantity. In predicting surface compositions, this ultimately limits the variety of compositions and segregation trends that can be observed and may induce size effects in small slabs. Some researchers have instead turned to machine learned potentials, which can be trained on DFT data and then utilized in larger simulations.<sup>[15]</sup> Although at first these potentials were narrowly limited to the specific data they were trained on, recent approaches have involved generating extremely large data sets that can be used to train generalized interatomic potentials that work across entire classes of materials.<sup>[16,17]</sup> Because these models have seen a wide variety of elemental atomic interactions, they can be fine-tuned on specific tasks or subdomains that they were not initially trained to describe, ultimately requiring less new data than a model trained from scratch.<sup>[4]</sup>

Of particular interest is the Open Catalyst Project (OCP), which developed a large data set and associated general interatomic potentials for the purposes of analyzing catalyst surfaces. It joins a growing number of large, high-quality DFT data sets for materials discovery; other notable databases include the Materials Project and Open DAC.<sup>[18,19]</sup> Besides big data, what unites these projects are community challenges – benchmark tasks that can be used to judge the effectiveness of machine learning algorithms for comprehending the data set as a whole.<sup>[20]</sup> Like ImageNet for computer vision and CASP for protein folding, these challenges provide a systematic framework for comparing the accuracy of various models used as DFT surrogates in materials science.<sup>[21,22]</sup> However, while models associated with the Open Catalyst Project continue to improve in their ability to predict DFT, there is no corresponding benchmark for accurately predicting catalyst activity, the ultimate goal of the project and the field of computational catalysis.<sup>[23]</sup>

The main reason that this benchmark does not exist is the difficulty of collecting large amounts of high-quality experimental catalysis data. While material properties or crystallized protein structures can be used to directly verify bulk formation

energy or protein structure predictions, the single crystal adsorption energy experiments corresponding to values that could be calculated with DFT or predicted by surrogate models are limited in scope and scale, and their connection to the activity of polycrystalline and nanoparticle catalysts is unclear. Meanwhile, attempting to benchmark DFT- and DFT surrogate-based catalyst discovery models using industrially relevant real-world examples is fraught, in part because the distribution of interface compositions and morphologies is extremely broad. Nanoparticle size, morphological dispersion, stability, and catalytic activity are sensitive to variations in synthesis and reactive environment that cannot be effectively described computationally with current methods and may be unreproducible experimentally.<sup>[10,24–26]</sup> Measurement of even conceptually simple properties like size distribution can be low-throughput and require the use of proxy metrics.<sup>[27]</sup>

The difficulty in establishing widely applicable data-driven models creates demand for physically motivated models like DFT surrogates that can describe trends in materials independently of real-world benchmarks. However, modern discovery campaigns that utilize both DFT and experiments tend to use DFT just to corroborate phenomena discovered by experimentation,<sup>[28]</sup> and purely computational campaigns can search broad regions of materials space but are difficult to connect to experimental outcomes.<sup>[29]</sup> To connect these two methods halfway, we propose comparing DFT surrogate models and experimental data on a simpler task that is still highly relevant to catalysis: understanding surface segregation in noble metal fcc ternary alloys.

### Noble Metal Alloys

Noble metal alloys are an important class of materials for studying surface segregation. Noble metals tend to be able to form stable nanoparticles and alloys, and have been investigated for a variety of industrial applications in catalysis and biomedicine.<sup>[30–33]</sup> Although noble metal- and especially platinum group-based catalysts are state of the art for many important reactions, their expense created a subfield of catalysis focused on reducing the use of precious metals by finely controlling surface composition and structure to maximize their precious metal surface exposure.<sup>[34,35]</sup> Earlier works also tried to design catalysts with strained overlayers to maximize activity, although in practice these materials often reconstructed.<sup>[36–38]</sup> Even in applications with near-stoichiometric quantities of noble metals, surface composition plays an important role in determining activity and selectivity.

Alloys of silver, gold, copper, palladium, and platinum have been found to have good catalytic properties for a wide range of reactions, including ethanol oxidation,<sup>[39]</sup> hydrogen storage,<sup>[40]</sup> and CO<sub>2</sub> reduction.<sup>[41]</sup> Mun et al. notes that, although pure copper offers good cost, activity, and stability for CO<sub>2</sub> reduction, it tends to result in the formation of a variety of gases that can be difficult to separate. Combining copper with palladium in an alloyed phase can reduce hydrogen adsorption and adsorbed CO protonation, improving Faradaic efficiency

and CO production.<sup>[41]</sup> Alternatively, combining gold with copper can stabilize copper against oxidation at higher temperatures, but the ratio of these elements must also be tuned to balance activity and stability.<sup>[42]</sup> Fu et al. discovered a similar phenomenon in AuPt alloys, where an underlayer of gold was found to stabilize the catalyst against poisoning by CO and organic compounds. Analysis of the surface combined with DFT study revealed that the formation of the catalytically active platinum overlayer depended on adsorbate-induced surface segregation.<sup>[43]</sup>

DFT has also been utilized to determine trends in the enthalpy of surface segregation in noble metal alloys. For example, Pt<sub>3</sub>M alloys were simulated using a  $2 \times 2 \times 5$  (111) slab with reasonable agreement between the energies of segregation calculated using DFT and the experimental surface energy of slabs of the same composition across a range of metals.<sup>[44]</sup> DFT was used to analyze the chemical and mechanical components of AuPd surface segregation in  $3 \times 3 \times 6$  and  $5 \times 5 \times 6$  slabs using an LDA functional<sup>[45]</sup> and general agreement with an experimental segregation isotherm at 800 K was found.

### Surface Segregation Simulations

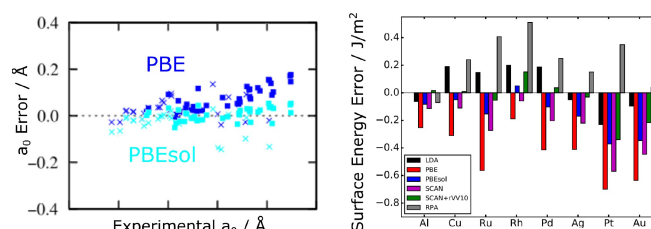
Monte Carlo simulations offer a method for incorporating entropic effects into segregation analyses and predicting equilibrium states.<sup>[46]</sup> However, size effects can confound slab simulations in smaller slabs, causing qualitatively incorrect behavior. If there are insufficient atoms of one type to coat the surface of a slab, an experimentally consistent equilibrium may be theoretically unobtainable. Segregation-driven depletion from the bulk may also substantially change the chemical environment of the interior of the slab, modifying chemical potentials and therefore predicted segregation relative to a larger slab.<sup>[47]</sup> Smaller slabs may also lack bulk-like interior atoms independent of segregation.<sup>[48]</sup>

Historically, lack of high-quality consistent segregation data across composition space has made direct evaluation of the effectiveness of Monte Carlo-derived surface segregation simulations challenging. Recently, composition spread alloy films (CSAFs) have investigated large regions of ternary alloy phase space with near-continuous resolution. Using a rotating mask to partially block the vapor flux from an electron beam deposition source, consistent multicomponent alloys with multiple composition gradients can be created in vacuum.<sup>[49]</sup> These CSAFs can then be analyzed using tools such as X-ray photoelectron spectroscopy (XPS), angle-resolved XPS, low-energy ion scattering (LEIS), and energy-dispersive X-ray spectroscopy (EDX) to generate continuous maps of surface and near-surface composition as a function of the bulk composition. This data offers a new lens into the efficacy of DFT-based Monte Carlo surface segregation studies.

A recent study by Yang et al. compared ternary AuCuPd composition spread surface segregation data with Monte Carlo simulations using energies from a DFT-based machine learned potential on (111)-oriented slabs.<sup>[50]</sup> They found that simulations

predicted some of the observed experimental behavior, but was qualitatively incorrect for others. The machine-learned potential used in that work was trained only on unrelaxed DFT single points of slabs oriented in the (111) direction, and so it could not be used on other surfaces. They conducted a preliminary investigation with DFT suggesting that surface relaxation may play an important role in (110)-oriented CuPd, and that the segregation energies on the (110) surface were consistent with experimentally observed segregation trends. A study of bimetallic segregation across many noble metals finds that segregation energies in bulk platinum are minimized along the (110) direction, although the orientation dependence in bulk palladium is nuanced.<sup>[51]</sup> The PBE functional used in both of these works is also less accurate than other more recent GGA functionals at predicting surface energies and other surface properties of common metals.<sup>[52]</sup> Although the PBE functional is commonly used across many applications in heterogeneous catalysis, the PBEsol functional is designed to better predict properties of dense solids and surfaces (see Figure 1).<sup>[53]</sup>

In this work, we develop a new data set of DFT relaxations across three disordered alloy systems – AgAuCu, AuCuPd, and CuPdPt – for the (111), (110), and (100) orientations using one of two DFT settings. Slabs ranging from 7 to 60 atoms were relaxed using either PBEsol functional with Becke-Johnson damping<sup>[54]</sup> ( $n = 2478$ ) or the PBE functional with no dispersion correction ( $N = 2118$ ). We fine-tune a model that had been pre-trained on the Open Catalyst data set to this data set and use this as the potential in canonical Monte Carlo simulations for each chemical system and orientation at a variety of bulk compositions ( $n \approx 200$  for each combination of orientation, system, and functional). This model architecture takes as its input a graph representation of a slab of arbitrary shape and size and predicts the total energy that DFT would calculate. See Section “Model Training” for further details on the model used. We examine trends in segregation as a function of slab composition and compare these to surface composition data from CSAF-based segregation experiments conducted on AgAuCu and AuCuPd alloys. Because the target used in training and validation, raw energy, is not directly relevant to Monte Carlo simulations, we analyze the relationship between accuracy in raw energy and transition probability, and we show how accuracy in both contexts varies as a function of training data. We vary the training data used to demonstrate how these



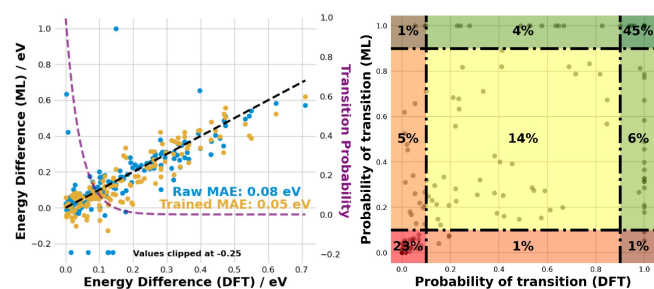
**Figure 1.** (left) Error in the lattice constants predicted using PBE (dark blue) and PBEsol (light blue). Figure taken directly from Zhang et al.<sup>[53]</sup> under the Creative Commons 3.0 Attribution License. (right) Error in the surface energies of common metals, including PBE (red) and PBEsol (blue). Taken directly from Patra et al.<sup>[52]</sup> under the PNAS License. Copyright 2017 National Academy of Sciences.

models can learn within a chemical system, across multiple systems, and how they can be fine-tuned from one category of DFT settings to another. Finally, we discuss how discrepancies between simulation and experiment can be understood primarily in terms of underlying errors in the DFT functionals used, highlighting the importance of functional accuracy in describing the behavior of real physical systems.

## Results and Discussion

### Model Accuracy and Learning Curves

We began by evaluating model accuracy in contexts where it extrapolates to slabs that are larger than its training data (see Section “Model Training” for methodological details). Figure 2 compares a model trained on small ( $\leq 35$ -atom) disordered AuCuPd slabs to a publicly available checkpoint that was pretrained only on RPBE slabs and no disordered structures.<sup>[56]</sup> Here, the test data is pairs of large (60-atom) disordered AuCuPd PBE slabs that differ by a single atomic exchange, mimicking a slab before and after a random Monte Carlo step. (Note that because the pretrained checkpoint was only trained on RPBE data, all outputs must be shifted by a constant amount). Although the  $> 1$  eV error of the pretrained benchmark would be unacceptably high in many contexts, accuracy is improved in a Monte Carlo context by cancellation of errors in the graph model that has been described before.<sup>[57]</sup> Because Monte Carlo simulations use only the energy differences between chemically and geometrically near-identical slabs, we can instead consider those differences to be the target: the differences in predicted energies of slabs that differ by only one atom swap were very close to the differences in DFT energies, with an error of only 0.08 eV using the pretrained model,

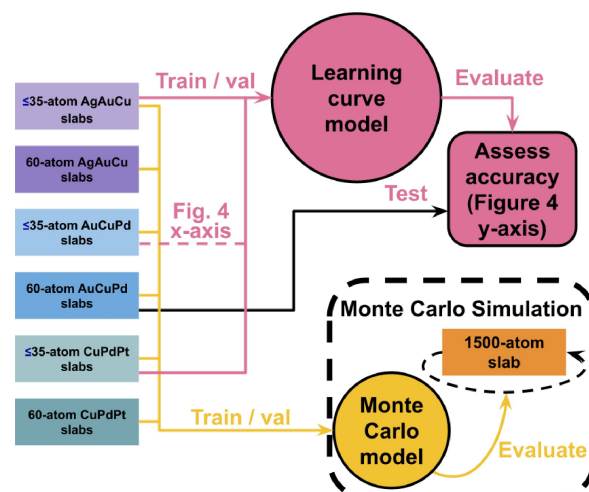


**Figure 2.** (left) The differences in energies between two slabs that have the positions of two atoms swapped between them calculated using DFT and the PBE functional (x-axis) and a machine learned model (y-axis) that is either a publicly available OCP checkpoint (blue) or one that has been fine-tuned on small AuCuPd slabs and all AgAuCu and AuCuPd slabs (blue). (right) The differences in PBE energies calculated using DFT (x-axis) and predicted by the fine-tuned model (y-axis) are transformed using a Boltzmann distribution at 500 K to compare the differences in probability of a transition occurring. The probabilities are classified into high ( $> 90\%$ ), medium ( $10\text{--}90\%$ ), and low ( $< 10\%$ ) probabilities, and a confusion matrix is overlaid on the parity plot. The mean-squared error of a probabilistic prediction is the Brier score. In this Figure, the Brier score is 0.05; with negative values excluded, this value rises to 0.07. See Figure 4 for further investigation of these metrics.

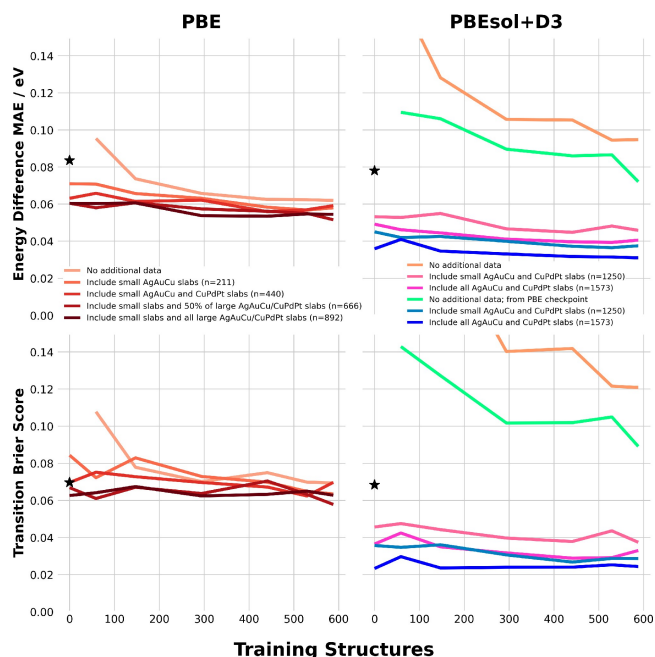
compared to 0.05 eV for the trained model. This translates into accurate probabilities in a Monte Carlo simulation.

Because this distribution is extremely modal, probabilities are shown in Figure 2 alongside a confusion matrix. This figure highlights that the trained model can accurately differentiate favorable and highly unfavorable swaps from moderately unfavorable swaps in a random alloy, but that regression accuracy within the indeterminate region is low. This region only includes about a fifth of the data set, however: half of all energy differences are negative and have transition probabilities of unity, and in about 60 % of the remaining energy swaps, energy differences are so high that the probability of a transition occurring is below 10 %. Accuracy in a probabilistic context can be determined by the Brier score (see Figure 4), which falls from 0.07 with the pretrained model to 0.05 with the fine-tuned one (or 0.09 and 0.07 respectively if the favorable, negative-energy swap is excluded from each pair). This finding indicates that while the fine-tuned models used in this work are accurate for the purposes of determining segregation, further training or higher-quality benchmarks would be needed to evaluate more complex phenomena such as long- or short-range order. This finding also shows that models trained on smaller slabs can extrapolate to larger ones, which is necessary because a relaxation of a 1500-atom slab using DFT would be intractable.

Figure 3 shows how the DFT calculations were used to develop the learning curves. The results of the learning curve experiments described in the Methods section are depicted in Figure 4. Two metrics are shown, corresponding to the metrics



**Figure 3.** A flowchart demonstrating how DFT data is used to construct one of the learning curves in Figure 4 and in Monte Carlo simulations, as described in Section “Model Training”. PBEsol + D3 DFT data generated during this study is shown on the left, and light-red flow lines demonstrate that the  $\leq 35$ -atom AgAuCu and CuPdPt slabs are pooled with variable amounts of  $\leq 35$ -atom AuCuPd slabs. This data is then split into 90 % training and 10 % validation and used to fine-tune various models. These models are then used to predict the energies of 60-atom AuCuPd slabs. Their accuracies are plotted in Figure 4 as the line of the same color, where the x-axis shows the number of  $\leq 35$ -atom AuCuPd slabs used in finetuning, and the y-axis shows the accuracy. Yellow flow lines show how the PBEsol + D3 Monte Carlo model is trained using all available data, and it is then used to iteratively update a 1500-atom slab as part of a Monte Carlo simulation.

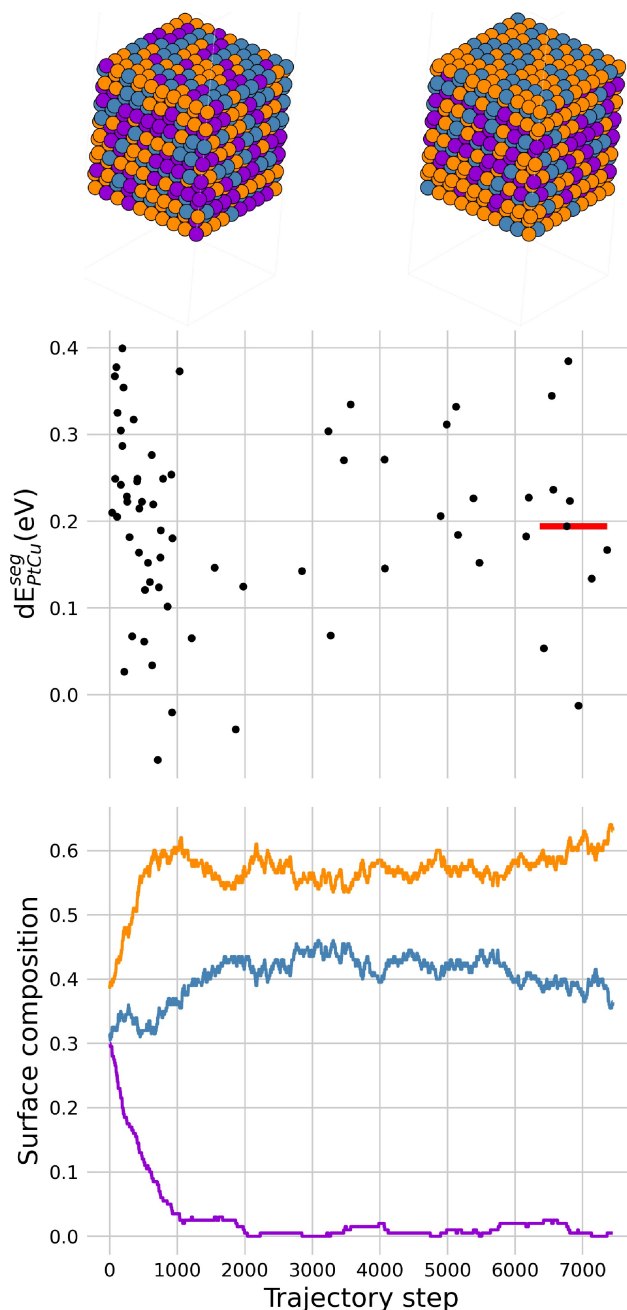


**Figure 4.** Learning curves for IS2RE models for PBE (left) and PBEsol + D3 (right) AuCuPd data. A number (indicated on the x axis) of small ( $\leq 35$ -atom) AuCuPd slabs are optionally combined with varying amounts of AgAuCu and CuPdPt data (indicated in the legend, with total the number of additional slabs) and split into 10% validation and 90% training. For all runs, the test data was 60-atom AuCuPd slabs. The top two plots compare the ability of models to predict the energy differences of pairs of slabs that had the positions of two atoms swapped, while the bottom two plots compare the differences between transition probabilities after transforming these energy differences with a Boltzmann distribution at 500 K. The accuracy of a publicly available checkpoint run in S2EF mode was included, represented by a black star.<sup>[56]</sup> Although this model does fairly well at predicting energy differences (MAE = 0.08 eV in both contexts) and probabilities (Brier score of 0.07 in both contexts), it is too slow at inference to be used in a Monte Carlo simulation. For further discussion, see the “Computational Methods” section.

described previously: the error in predicting the energy difference between pairs of atom-exchanged slabs and the Brier score in predicting the probabilities of transition based on applying a Boltzmann distribution to energy differences. Consistent with the previous section, the pretrained model is highly accurate. The authors note that models can be trained in S2EF mode to yield significantly higher accuracies (0.03 eV error in predicting energy differences) at significant cost (approximately 10 $\times$  increase) to simulation times. Among the trained IS2RE models, those trained on PBE data converged relatively quickly, needing only about 200 data points in either the AuCuPd or AgAuCu system to reach the level of the pretrained S2EF model, and then plateauing. Models trained on the dispersion-corrected PBEsol data required significantly more data to reach saturation ( $n \approx 1000$ ), but pretraining on all PBE data ( $n = 1872$ ) results in small but noticeable improvements in learning. Like the PBE training curves, significant improvements were also found by training on data from chemically similar systems: in fact, if AgAuCu and CuPdPt data is excluded, there is insufficient AuCuPd PBEsol data ( $n \approx 600$ ) to reach data saturation. With these data sets included, models are highly accurate even without AuCuPd data, implying that this method

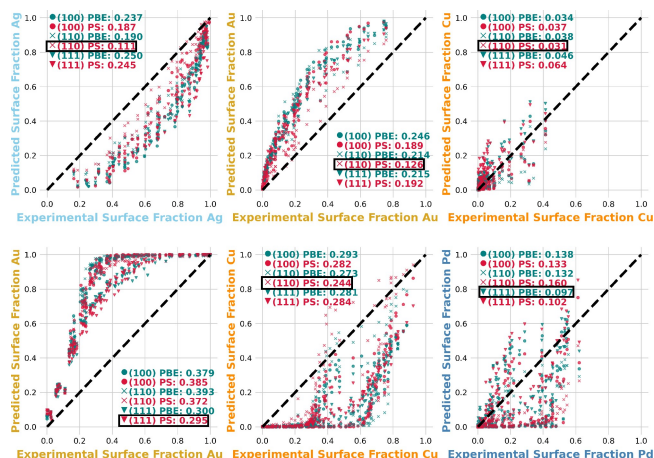
may be extended to similar systems with little or no additional training data.

As shown in Figure 5, there was significant variation in atomic segregation energies (defined in the “Computational Methods” section) within chemically similar slabs. These values



**Figure 5.** A Monte Carlo simulation progressing from left to right. (top) The slab before (left) and after (right) simulation. (middle) Observed Cu–Pt segregation events, where one atom was present in the top layer and another was not, and their associated energy differences. The average steady-state segregation energy is marked in red: this value is plotted as a function of composition in Figure 9. (bottom) The overall surface composition over the course of the simulation. The surface composition is initially approximately identical to the slab, and over time, copper segregates over platinum. Cu is in orange, Pd in blue, and Pt in purple. The final surface composition is compared to the slab composition using Eq. (2), defined in Section “Monte Carlo Simulations”.

also changed as slabs segregated, and they differed systematically between the two functionals. However, because there are 13 interior layers and 2 surface layers, most attempted Monte Carlo steps did not involve surface segregation. This effect along with surface depletion resulted in some regions of ternary composition space having few or no observed segregation events at equilibrium during the sampling window.

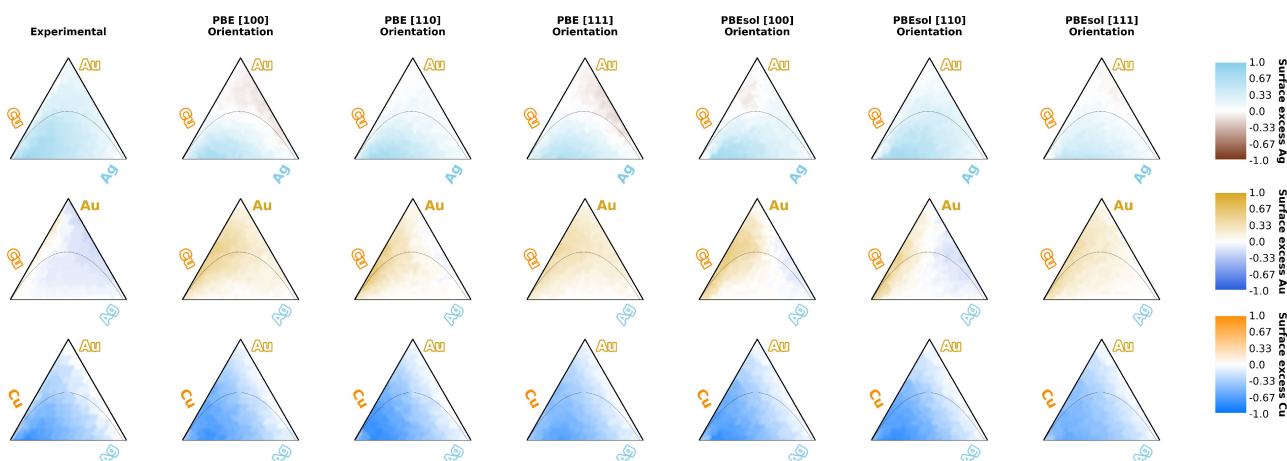


**Figure 6.** (top) Parity plots for predictions of experimental CSAF-derived measurements of surface composition of silver (left), gold (middle), and copper (right) for the AgAuCu ternary system for PBE (teal) and PBEsol + D3 (red) using slabs oriented in the 100 (circle), (110) (x), and (111) (triangle) directions. The mean absolute errors across all experimental data points for each setting are listed and the best setting for each plot is highlighted with a black box. A black dashed line is drawn for parity. (bottom) Parity plots for predictions of experimental CSAF-derived measurements of surface composition of gold (left), copper (middle), and palladium (right) for the AuCuPd ternary system for PBE (teal) and PBEsol + D3 (red) using slabs oriented in the (100) (circle), (110) (x), and (111) (triangle) directions. The mean absolute errors across all experimental data points for each setting are listed. For parity plots of thermodynamic segregation energies, see the Supporting Information.

## Experimental Comparison

Parity plots comparing simulated and experimentally observed segregation trends across both exchange-correlation functionals and in all orientations are shown in Figure 6. Simulated surface compositions were linearly interpolated onto compositions matching experimental observations. In describing the AgAuCu ternary space, simulations that utilized surrogates for the PBEsol exchange-correlation functional with D3 dispersion corrections and slabs oriented in the (110) direction more closely matched experimentally observed trends in surface composition than any other setting. We note, however, that we do not know the relative fraction of (110) surfaces in the polycrystalline samples. No setting accurately predicted the observed strong silver segregation, although surface silver mobility was observed experimentally and was not included as a factor in simulations. Experimentally observed segregation in the AuCuPd ternary system is predicted less well: gold is predicted to segregate over copper and palladium much more strongly than was observed in experiment, and copper segregated into the bulk much more strongly in simulated slabs than in CSAFs. Although no setting agreed better with the experimental surface segregation trends observed on AuCuPd CSAFs, simulations utilizing slabs oriented in the (110) and (111) directions were more effective than those that used slabs oriented in the (100) direction, and simulations utilizing a surrogate for the PBEsol functional were at least as good as those utilizing the PBE surrogate. However, these parity plots do not capture composition-level trends, which will be examined next.

The effects of variation in predicted surface composition as a function of bulk composition at varying simulation exchange-correlation functionals and surface orientations are shown for the AgAuCu ternary system in Figure 7. A comparison of the experimental data with simulations suggests that the best agreement comes from slabs in the (110) orientation and the

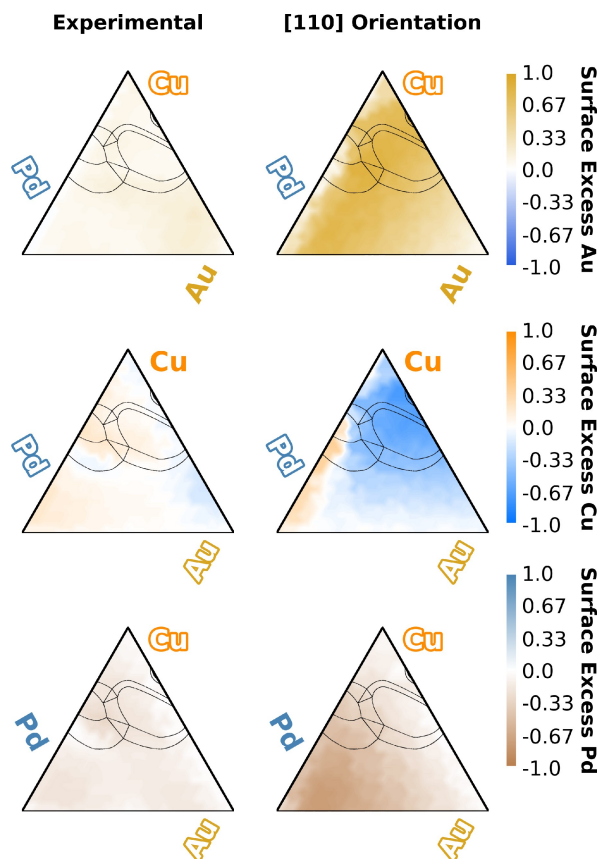


**Figure 7.** A comparison between experimental ternary segregation enrichment (left-most column) and surface excess calculated using Eq. (2) (see Section "Monte Carlo Simulations" and Figure 5 for definitions) on Monte Carlo simulations for the AgAuCu ternary, both at 800 K. Surface excess compositions of silver (top), gold (middle), and copper (bottom) are compared between experiment (column 1) and the results of simulations on the (100) (columns 2 and 5), (110) (columns 3 and 6), and (111) (columns 4 and 7) for PBE (columns 2–4) and PBEsol + D3 (columns 5–7). An experimental ternary phase diagram for AgAuCu at 798 K is overlaid on each plot, where the fcc alloy is present at all corners of the ternary, and the region along the Ag–Cu binary exhibits Ag–Cu immiscibility, separating into Ag-rich and Cu-rich fcc alloys.<sup>[58]</sup>

model simulating the dispersion-corrected PBEsol functional. This set of simulations accurately predicts silver segregation throughout the ternary region and surface gold enrichment along the Au–Cu binary and depletion along the Ag–Au binary. The (110) as suggested by Yang et al. is the orientation that best captures segregation trends in the CuPd system,<sup>[50]</sup> in spite of the fact that the (111) orientation of fcc crystals is typically the most stable. This behavior may be explained by Han et al.,<sup>[59]</sup> which finds that a significant mismatch in atomic size can drive segregation especially to step and kink sites where undercoordination can relieve lattice strain. Gold and silver are both larger than copper and have lower surface energies, suggesting that multicomponent alloys may prefer the (110) orientation to minimize lattice strain. This is supported by other works finding alloying is most favored in the (110) orientation for the Cu–Pd binary and unfavorable in the close-packed (111) direction<sup>[60]</sup> and that the (110) orientation minimized interfacial energy in ordering Ag–Au and Au–Cu systems.<sup>[61]</sup> An additional factor complicating predictions on AgAuCu is the high lateral surface mobility of Ag, which may be mediated by the unique structure of the (110) orientation.<sup>[62]</sup> This could explain why the degree of silver segregation is systematically underpredicted. Although the surface structure and orientation of experimental alloys is not known, they probably consist of polycrystalline grains and the surface is likely dominated by a distribution of low Miller index orientations.

A similar trend was found for AuCuPd, where simulations run using (110)-oriented slabs and a model trained on data generated using the D3-corrected PBEsol functional were in best agreement with the experimental trends. Figure 8 compares experimental data to simulations for only these settings; see the Supporting Information for equivalent plots for results across all orientations and functionals. An experimental phase diagram at 623 K is overlaid on each plot: all interior phases are tetrahedral, whereas both phases present at the ternary points are fcc.<sup>[58]</sup> However, comparing only the fcc regions, comparison with experiment makes clear that the surface favorability of gold is being overpredicted, a result of the surface energy of pure gold being underpredicted by both the PBE and PBEsol functionals (see Figure 1). Surface energy errors calculated using PBEsol are significantly reduced relative to PBE, resulting in improved segregation predictions, but the surface energies predicted for silver, gold, and platinum with this functional are still substantially lower than those measured by experiment.

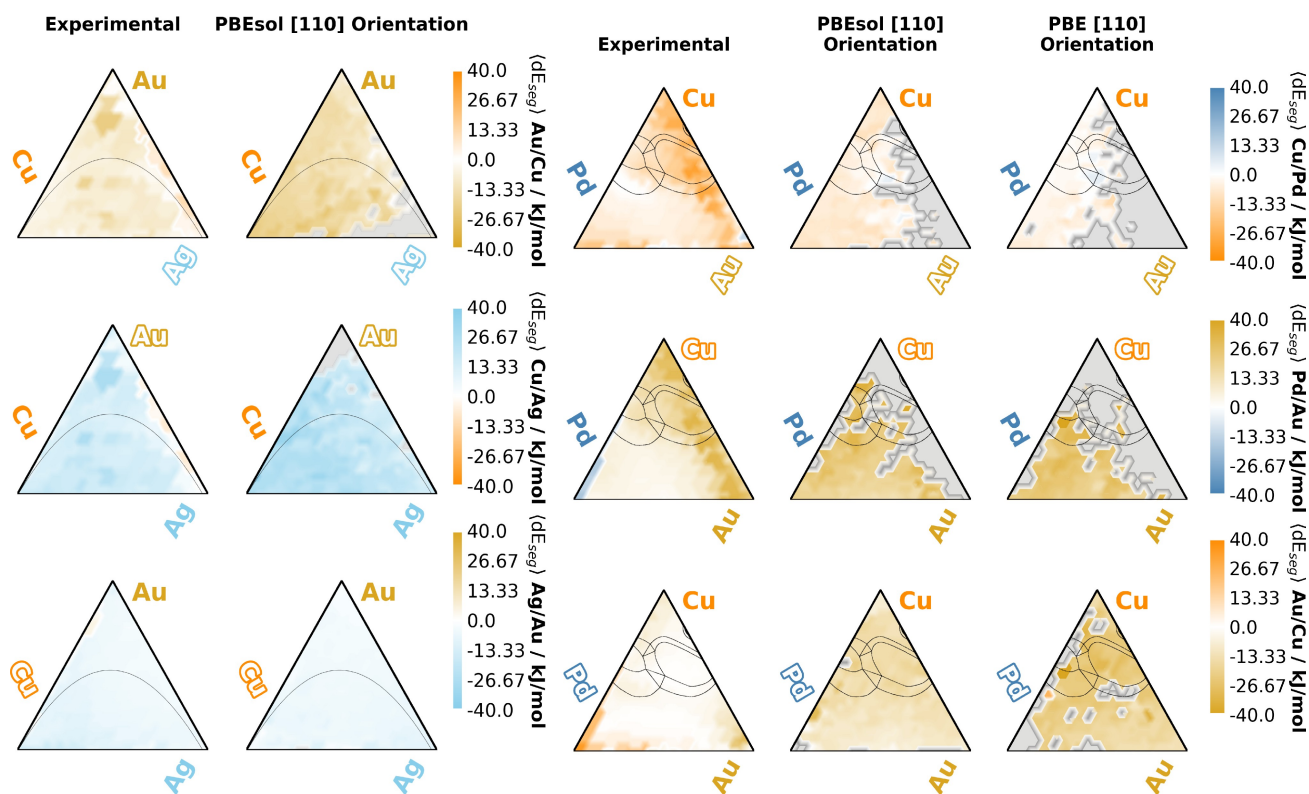
Although DFT energies of segregation calculated using Eq. (4) and macroscopic free energies of segregation calculated with Eq. (5) are not directly analogous, comparing them is a useful exercise for two reasons. (These equations are defined in Section “Segregation Energies”). In cases where there is extreme enrichment or depletion of one component, the energetics of swapping the other two elements can give additional computational insight into chemical equilibria where the surface excess compositions of these components are strongly determined by the third component. In these cases, size effects can distort binary equilibria in model slabs. Additionally, observations of trends in  $\langle \Delta E_{AB}^{\text{seg}} \rangle$  (calculated using Eq. (4)) within a ternary can



**Figure 8.** (left) Experimental AuCuPd segregation data and (right) PBEsol-based Monte Carlo simulations on (110) oriented slabs for each component of the AuCuPd ternary system, both at 500 K. The experimental ternary phase diagram at 623 K is overlaid on the plot: the phase at all three corners is soluble fcc; the other phases are ternary but not fcc.<sup>[58]</sup>

give additional insight into the sources of distortion between experiment and theory.

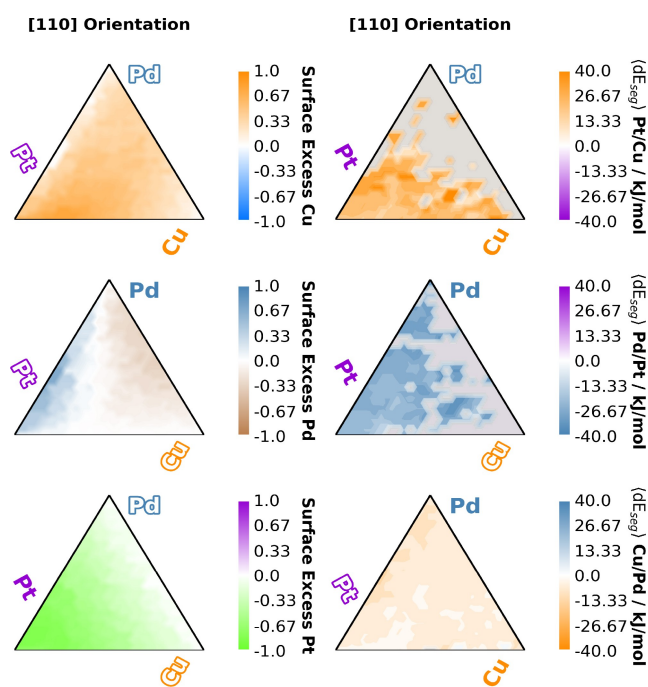
Figure 9 provides further support to the idea that differences between theory and experiment can be explained by errors in surface energies. In the AgAuCu ternary system, the average segregation energy of the Ag–Au binary somewhat favors silver, correctly predicting that silver segregates slightly over gold across the ternary space; the surface energies of silver and gold are underpredicted by approximately the same amount. Meanwhile, the average segregation energies of both silver and gold over copper at simulated equilibrium are higher than those inferred from observed experimental trends. In AuCuPd, a parallel story can be told: although the substantial underprediction of gold's surface energy results in more gold segregation than observed, simulations predict favorable copper segregation over palladium, consistent with the relatively smaller and more consistent underpredictions for those elements. Comparing average segregation energies between PBE-based and PBEsol-based simulations in the fcc regions along the AuCuPd Au–Pd binary and near pure Cu, the PBE results showed significantly higher underprediction of the gold surface energy resulting in greater overestimation of the average gold surface segregation energy. Additionally, copper



**Figure 9.** Experimental  $\Delta G_{\text{seg}}$  (columns 1 and 3 respectively) calculated using Eq. (5) plotted next to the average DFT energy of segregation (columns 2 and 4) calculated with Eq. (4) using PBEsol-based Monte Carlo simulations on (110)-oriented surfaces for the 800 K AgAuCu (left) and 500 K AuCuPd (right) ternaries, overlaid with experimental ternary phase diagrams at 798 K and 623 K respectively.<sup>[58]</sup> Plots are grey in places where no segregation event was observed at equilibrium. The results of PBE-based simulations of AuCuPd slabs on (110) are also included for comparison (column 5).

segregation over palladium is better predicted by PBEsol segregation energies than PBE, possibly owing to substantially higher and more varied errors in the surface energies of those two elements with the PBE functional.

Figure 10 shows predictions for the (110) orientation of the CuPdPt ternary system at 600 K using the PBEsol-based model. This choice of system is constrained by the analytical methods used. Computationally, the Monte Carlo algorithm would be significantly hindered by the need to infer magnetic states, precluding analysis of iron, cobalt, or nickel ternaries. Additionally, elements from different rows of the periodic table are suitable for comparison with ion scattering experiments, the spectra of which are well resolved for elements of different masses. Because these elements are likely to have very different lattice constants, this selection could bias the observations made here in favor of the (110) orientation, which can reduce lattice strain relative to the closer-packed (111) and (100) due to its more open rows. However, the CuPdPt system is unlikely to violate this trend. In keeping with the comparisons made using the AgAuCu and AuCuPd ternaries and surface energy errors shown in Figure 1, we expect the Cu–Pd binary to be well-predicted by this method, but platinum to be overrepresented on the surface. As such, it is interesting that even this model predicts platinum to strongly deplete from the surface across the ternary space.



**Figure 10.** Monte Carlo predicted segregation (left) and average predicted segregation energy (right) for the CuPdPt ternary system at 600 K using a PBEsol-based model and (110)-oriented model surfaces. Segregation energies are gray when few segregation events were observed at equilibrium.

## Conclusions

In this work, we simulated surface segregation using a lattice Monte Carlo algorithm for three ternary systems, AgAuCu, AuCuPd, and CuPdPt, and three crystallographic orientations, (111), (110), and (100), using  $10 \times 10 \times 15$  simulated fcc slabs. The potential used to evaluate slab energies is a graph neural network trained on fcc slabs up to 60 atoms in size in each ternary system and orientation that were relaxed using the PBEsol functional with D3 corrections. We first demonstrated that this model can extrapolate to larger slabs at an acceptable accuracy for Monte Carlo simulations. Then, we evaluated its ability to make predictions across chemical composition space and its data efficiency and found that it is data saturated and capable of quickly learning a new ternary given that it has seen each of the elements in the new composition. After running 200 Monte Carlo simulations for each ternary system and orientation, we compared sets of simulations to experimental data obtained for AgAuCu and AuCuPd segregation from other studies. We found that the (110) orientations showed the best agreement with experimental segregation trends for both ternaries evaluated, and that simulations using a model that simulates relaxation using the PBEsol functional with D3 corrections is more accurate than using one trained on PBE data. We then examined atomic segregation energy to compare binary segregation trends against experimental data where one component dominates the simulated surface. In both aggregate and atomic simulated segregation, we found that differences between DFT and experiment can largely be explained by known biases in the choice of functional, highlighting the importance of accurate functionals for materials discovery efforts and the need for improvements in simulation technologies to achieve high predictive accuracies in furtherance of those efforts. This work comprises experimental validation of key components of computational materials interface discovery efforts at large scales and high resolution across a span of catalytically important compositions, highlighting areas where methodological improvements are needed and demonstrating the potential of computational simulations to aid in materials discovery efforts.

## Computational Methods

### DFT Settings

The Vienna Ab Initio Simulation Package (VASP) with Projector Augmented Wave pseudopotentials was used to conduct DFT atomic relaxations of simulated ternary alloy slabs.<sup>[63,64]</sup> Two functionals were used: the GGA-PBEsol<sup>[53]</sup> with D3 Beck-Johnson damping<sup>[54]</sup> and GGA-PBE.<sup>[65]</sup> A kinetic energy cutoff of 400 eV was used, and KPOINTS were set to a minimum spacing of  $0.2 \text{ \AA}^{-1}$ . In the ionic relaxation loop, only atomic positions were allowed to relax, and they were optimized using the conjugate gradient algorithm until the total energy difference fell below 0.0001 eV or to 150 atomic steps. Slabs that did not converge were discarded. Computational resources used in this work are discussed in the Supporting Information.

### DFT Slab Creation

Slabs of various sizes from 7–35 atoms oriented in the (100), (110), or (111) direction were created using ASE and Pymatgen.<sup>[18,66]</sup> The slabs were initialized on an ideal fcc lattice with the lattice constants set randomly between those of the smallest and largest elements present. Slab compositions were assigned using a Dirichlet distribution. Elements were randomly assigned to lattice sites so that slabs were either ternary AgAuCu, AuCuPd, or CuPdPt, or a constituent binary. To test the ability of interatomic potentials to evaluate energetic differences in larger slabs, 60-atom ternary AgAuCu, AuCuPd, or CuPdPt slabs were created in pairs, with the position of two atoms swapped between them and the lattice constant set according to Vegard's Law using experimentally observed lattice constants.

### Model Training

For the models used in Monte Carlo simulations, all slabs relaxed with a given pseudopotential were used to fine-tune a message-passing neural network using the GemNet-OC architecture starting from a publicly available checkpoint trained on the OC20 and OC22 data sets.<sup>[56,67,68]</sup> To assess model learning, various other models were trained starting from the same checkpoint or the checkpoint trained on all PBE data. To construct learning curves, either 0%, 10%, 25%, 50%, 75%, 90%, or 100% of all AuCuPd slabs less than 60 atoms for either functional was used. This data was then optionally combined with other subsets of the remaining data. Three additional learning curves were created using the slabs relaxed with PBE: one included in its training data all AgAuCu slabs under 60 atoms in its training data, another included all AgAuCu and CuPdPt training slabs under 60 atoms, and a third used all AgAuCu and CuPdPt slabs. With the PBEsol dispersion-corrected data, one learning curve was run using only AuCuPd data, another included all AgAuCu and CuPdPt slabs, and two more learning curves were developed identical to these except that they were initialized from a checkpoint that was trained on all of the PBE data. In all cases, 10% of the data available to the model for training was reserved for validation.

### Monte Carlo Simulations

Monte Carlo simulations were run for three ternary systems: Ag–Au–Cu at 800 K, Au–Cu–Pd at 500 K, and Cu–Pd–Pt at 600 K. All Monte Carlo simulations were initialized using a unique random number generator (RNG) instance that was stored and loaded alongside the simulation. Ternary AgAuCu, AuCuPd, and CuPdPt compositions created using a Dirichlet distribution were initialized randomly onto a  $10 \times 10 \times 15$  (100), (110), or (111) slab in a perfect fcc lattice with the lattice constant set using Vegard's law. At every Monte Carlo step, a random pair of atoms was swapped according to the RNG instance, atomic positions were reinitialized onto an ideal lattice, their relaxed energies were predicted using one of the interatomic potentials described in the previous section, and the energies calculated by the interatomic potential were used to determine whether a swap occurred according to a Boltzmann distribution defined by Eq. (1) that was controlled by the RNG instance.

$$P_{\text{trans}}(\Delta E^{\text{swap}}) = \begin{cases} 1 & \text{if } \Delta E^{\text{swap}} < 0 \\ \exp\left(\frac{-\Delta E^{\text{swap}}}{RT}\right) & \text{otherwise} \end{cases} \quad (1)$$

In the Monte Carlo simulations, the relaxed energy must be predicted directly, because relaxing a slab using an interatomic

potential and then predicting the energy takes an order of magnitude longer (0.6 seconds vs 6 seconds on an internal benchmark). However, relaxations using a pretrained interatomic potential were also included for reference. The authors note that with relaxation-based simulations, if atomic positions were not reinitialized to an ideal lattice, the simulation became strongly biased against making atomic swaps and quickly became stuck in local minima. All simulations were determined to be converged under any of the following criteria: at any point, if the number of consecutive Boltzmann steps taken was above 20, or 250 atomic swaps had been attempted without finding two different elements, or 15000 atomic swaps had been made, or 2000 steps had been taken without a new energy minimum being observed. The surface composition was determined using the 100 surface atoms on either side of the slab. The surface excess of any component can then be calculated using Eq. (2):

$$x_i^{\text{excess}} = N_i^{\text{surf}}/200 - N_i^{\text{slab}}/1500 \quad (2)$$

### Segregation Energies

The energy of segregation was calculated directly from the DFT energies of slabs that differed by a single atom swap where that swap involving a surface atom and a bulk atom. This is defined in Eq. (3), where  $E[A_{\text{surf}}, B_{\text{bulk}}]$  is the energy of the slab in which element A is on the surface and element B is in the bulk.

$$\Delta E_{AB}^{\text{seg}} = E[A_{\text{surf}}, B_{\text{bulk}}] - E[A_{\text{bulk}}, B_{\text{surf}}] \quad (3)$$

To define an average segregation energy, these values were then averaged over later steps of the simulation that involved a swap between a surface atom and a bulk atom, resulting in an expectation value for  $\Delta E_{AB}^{\text{seg}}$  shown in Eq. (4) that has been weighted according to the Boltzmann sampling used to construct the sequence.

$$\langle \Delta E_{AB}^{\text{seg}} \rangle = \frac{1}{n} \sum_{i=1}^n \Delta E_{AB}^{\text{seg}}(i) P_{\text{trans}}(\Delta E_{AB}^{\text{seg}}(i)) \quad (4)$$

It is important to note that, while both equations are antisymmetric relative to atomic exchange ( $\Delta E_{AB} = -\Delta E_{BA}$  and  $\langle \Delta E_{AB} \rangle = \langle -\Delta E_{BA} \rangle$ ), and the atomic reaction cycle is thermodynamically neutral ( $\Delta E_{A_1 B_2}^{\text{seg}} + \Delta E_{B_1 C_2}^{\text{seg}} + \Delta E_{C_1 A_2}^{\text{seg}} = 0$ ) for any two sites 1 and 2 and components A, B, C, the analogous statistical cycle does not hold ( $\langle \Delta E_{AB}^{\text{seg}} \rangle + \langle \Delta E_{BC}^{\text{seg}} \rangle + \langle \Delta E_{CA}^{\text{seg}} \rangle \neq 0$ ). Using experimental measurements of surface and bulk composition, an empirical free energy of segregation can also be calculated using Eq. (5), where  $A_{\text{surf}}$  is the composition of component A measured in the surface of the slab.

$$\Delta G_{AB}^{\text{seg}} = -RT \ln \frac{|A_{\text{bulk}}||B_{\text{surf}}|}{|A_{\text{surf}}||B_{\text{bulk}}|} \quad (5)$$

### Acknowledgements

This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award BES-ERCAPm2755. This work was supported by NSF DMREF Award CBET-1921946.

### Conflict of Interests

The authors declare no conflict of interest.

### Data Availability Statement

DFT calculations that support the findings of this study are openly available in Zenodo at <https://doi.org/10.5281/zenodo.10501343>.

**Keywords:** Monte Carlo · surface chemistry · density functional calculations · surface segregation · machine learning

- [1] N. K. Katiyar, K. Biswas, J.-W. Yeh, S. Sharma, C. S. Tiwary, *Nano Energy* **2021**, *88*, 106261.
- [2] J. K. Nørskov, M. Scheffler, H. Toulhoat, *MRS Bull.* **2006**, *31*, 669.
- [3] O. Isayev, D. Fourches, E. N. Muratov, C. Oses, K. Rasch, A. Tropsha, S. Curtarolo, *Chem. Mater.* **2015**, *27*, 735.
- [4] N. Shoghi, A. Kolluru, J. R. Kitchin, Z. W. Ulissi, C. L. Zitnick, B. M. Wood, *arXiv preprint arXiv:2310.16802* **2023**.
- [5] S. B. Torrisi, M. Z. Bazant, A. E. Cohen, M. G. Cho, J. S. Hummelshøj, L. Hung, G. Kamat, A. Khajeh, A. Kolluru, X. Lei, et al., *APL Machine Learning* **2023**, *1*.
- [6] J. Leeman, Y. Liu, J. Stiles, S. Lee, P. Bhatt, L. Schoop, R. Palgrave, *chemrxiv - 10.26434/chemrxiv-2024-5p9j4* **2024**.
- [7] S. Sun, H. Li, Z. J. Xu, *Joule* **2018**, *2*, 1024.
- [8] G. Prieto, J. Zečević, H. Friedrich, K. P. De Jong, P. E. De Jongh, *Nat. Mater.* **2013**, *12*, 34.
- [9] R. Narayanan, M. A. El-Sayed, *J. Phys. Chem. B* **2005**, *109*, 12663.
- [10] K. Broderick, E. Lopato, B. Wander, S. Bernhard, J. Kitchin, Z. Ulissi, *Appl. Catal. B* **2023**, *320*, 121959.
- [11] K. J. Andersson, F. Calle-Vallejo, J. Rossmeisl, I. Chorkendorff, *J. Am. Chem. Soc.* **2009**, *131*, 2404.
- [12] A. A. Vega, R. C. Newman, *J. Electrochem. Soc.* **2013**, *161*, C11.
- [13] Z.-J. Wang, J. Chen, Y. Huang, Z.-X. Chen, *Materials Today Communications* **2021**, *28*, 102475.
- [14] K. G. Papanikolaou, M. T. Darby, M. Stamatakis, *J. Phys. Chem. C* **2019**, *123*, 9128.
- [15] M. Liu, Y. Yang, J. R. Kitchin, *J. Chem. Phys.* **2021**, *154*, 134701.
- [16] V. L. Deringer, M. A. Caro, G. Csányi, *Adv. Mater.* **2019**, *31*, 1902765.
- [17] J. Riebesell, R. E. Goodall, A. Jain, P. Benner, K. A. Persson, A. A. Lee, *arXiv preprint arXiv:2308.14920* **2023**.
- [18] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, et al., *APL Mater.* **2013**, *1*.
- [19] A. Sriram, S. Choi, X. Yu, L. M. Brabson, A. Das, Z. Ulissi, M. Uyttendaele, A. J. Medford, D. S. Sholl, *arXiv preprint arXiv:2311.00341* **2023**.
- [20] A. Dunn, Q. Wang, A. Ganose, D. Dopp, A. Jain, *npj Computational Materials* **2020**, *6*, 138.
- [21] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: A large-scale hierarchical image database, in *2009 IEEE conference on computer vision and pattern recognition*, IEEE **2009** pages 248–255.
- [22] A. Kryshchuk, T. Schwede, M. Topf, K. Fidelis, J. Moul, *Proteins Struct. Funct. Bioinf.* **2019**, *87*, 1011.
- [23] Y.-L. Liao, B. Wood, A. Das, T. Smidt, *arXiv preprint arXiv:2306.12059* **2023**.
- [24] Z. C. Simon, E. M. Lopato, M. Bhat, P. J. Moncure, S. M. Bernhard, J. R. Kitchin, S. Bernhard, J. E. Millstone, *ChemCatChem* **2022**, *14*, e202101551.
- [25] P. G. Jamkhande, N. W. Ghule, A. H. Bamer, M. G. Kalaskar, *J. Drug Delivery Sci. Technol.* **2019**, *53*, 101174.
- [26] R. Han, K. S. Walton, D. S. Sholl, *Annu. Rev. Chem. Biomol. Eng.* **2019**, *10*, 43.
- [27] M. M. Modena, B. Rühle, T. P. Burg, S. Wuttke, *Adv. Mater.* **2019**, *31*, 1901556.
- [28] A. Kulkarni, S. Siahrostami, A. Patel, J. K. Nørskov, *Chem. Rev.* **2018**, *118*, 2302.
- [29] K. Tran, W. Neiswanger, K. Broderick, E. Xing, J. Schneider, Z. W. Ulissi, *J. Chem. Phys.* **2021**, *154*.
- [30] K. Loza, M. Heggen, M. Eppel, *Adv. Funct. Mater.* **2020**, *30*, 1909260.

- [31] S. Sohn, Y. Liu, J. Liu, P. Gong, S. Prades-Rodel, A. Blatter, B. E. Scanley, C. C. Broadbridge, J. Schroers, *Scr. Mater.* **2017**, *126*, 29.
- [32] A. Zaleska-Medynska, M. Marchelek, M. Diak, E. Grabowska, *Adv. Colloid Interface Sci.* **2016**, *229*, 80.
- [33] G. Habibullah, J. Viktorova, T. Ruml, *Nanoscale Res. Lett.* **2021**, *16*, 1.
- [34] N. Cheng, L. Zhang, K. Doyle-Davis, X. Sun, *Electrochem. Energy Rev.* **2019**, *2*, 539.
- [35] S. Gong, Y.-X. Zhang, Z. Niu, *ACS Catal.* **2020**, *10*, 10886.
- [36] V. Stamenkovic, B. S. Mun, K. J. Mayrhofer, P. N. Ross, N. M. Markovic, J. Rossmeisl, J. Greeley, J. K. Nørskov, *Angew. Chem. Int. Ed.* **2006**, *45*, 2897.
- [37] J. Greeley, T. F. Jaramillo, J. Bonde, I. Chorkendorff, J. K. Nørskov, *Nat. Mater.* **2006**, *5*, 909.
- [38] J. Greeley, J. K. Nørskov, L. A. Kibler, A. M. El-Aziz, D. M. Kolb, *ChemPhysChem* **2006**, *7*, 1032.
- [39] K.-U. Lee, J. Y. Byun, H.-J. Shin, S. H. Kim, *J. Alloys Compd.* **2020**, *842*, 155847.
- [40] K. Nakajima, M. Tominaga, M. Waseda, H. Miura, T. Shishido, *ACS Sustainable Chem. Eng.* **2019**, *7*, 6522.
- [41] Y. Mun, S. Lee, A. Cho, S. Kim, J. W. Han, J. Lee, *Appl. Catal. B* **2019**, *246*, 82.
- [42] Z. Xu, E. Lai, Y. Shao-Horn, K. Hamad-Schifferli, *Chem. Commun.* **2012**, *48*, 5626.
- [43] X. Fu, J. B. Pedersen, Y. Zhou, M. Saccoccio, S. Li, R. Sažinas, K. Li, S. Z. Andersen, A. Xu, N. H. Deissler, et al., *Science* **2023**, *379*, 707.
- [44] Y. Ma, P. B. Balbuena, *Surf. Sci.* **2008**, *602*, 107.
- [45] J. Creuze, H. Guesmi, C. Mottet, B. Zhu, B. Legrand, *Surf. Sci.* **2015**, *639*, 48.
- [46] M. Polak, L. Rubinovich, *Surf. Sci. Rep.* **2000**, *38*, 127.
- [47] C. Cserhati, I. Szabo, D. Beke, *J. Appl. Phys.* **1998**, *83*, 3021.
- [48] P. A. Schultz, *Phys. Rev. B* **2021**, *103*, 195426.
- [49] B. Fleutot, J. B. Miller, A. J. Gellman, *J. Vac. Sci. Technol. A* **2012**, *30*.
- [50] Y. Yang, Z. Guo, A. J. Gellman, J. R. Kitchin, *J. Phys. Chem. C* **2022**, *126*, 1800.
- [51] L. Farsi, N. A. Deskins, *Phys. Chem. Chem. Phys.* **2019**, *21*, 23626.
- [52] A. Patra, J. E. Bates, J. Sun, J. P. Perdew, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E9188.
- [53] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, K. Burke, *Phys. Rev. Lett.* **2008**, *100*, 136406.
- [54] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456.
- [55] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, M. Scheffler, *New J. Phys.* **2018**, *20*, 063020.
- [56] J. Gasteiger, M. Shuaibi, A. Sriram, S. Günnemann, Z. Ulissi, C. L. Zitnick, A. Das, *arXiv preprint arXiv:2204.02782* **2022**.
- [57] J. Ock, T. Tian, J. Kitchin, Z. Ulissi, *J. Chem. Phys.* **2023**, *158*.
- [58] Au (Gold) Ternary Alloy Phase Diagrams, in *Alloy Phase Diagrams*, ASM International **2016**.
- [59] J. W. Han, J. R. Kitchin, D. S. Sholl, *J. Chem. Phys.* **2009**, *130*.
- [60] L. Pielsticker, I. Zegkinoglou, Z.-K. Han, J. J. Navarro, S. Kunze, O. Karslioglu, S. V. Levchenko, B. Roldan Cuenya, *J. Phys. Chem. Lett.* **2021**, *12*, 2570.
- [61] V. Ozoliņš, C. Wolverton, A. Zunger, *Phys. Rev. B* **1998**, *57*, 4816.
- [62] J. Paez-Ornelas, H. Fernández-Escamilla, N. Takeuchi, J. Guerrero-Sánchez, *Materials Today Communications* **2020**, *25*, 101461.
- [63] G. Kresse, J. Furthmüller, *Phys. Rev. B* **1996**, *54*, 11169.
- [64] G. Kresse, D. Joubert, *Phys. Rev. B* **1999**, *59*, 1758.
- [65] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865.
- [66] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dulák, J. Friis, M. N. Groves, B. Hammer, C. Hargus, et al., *J. Phys. Condens. Matter* **2017**, *29*, 273002.
- [67] L. Chanussot, A. Das, S. Goyal, T. Lavril, M. Shuaibi, M. Riviere, K. Tran, J. Heras-Domingo, C. Ho, W. Hu, et al., *ACS Catal.* **2021**, *11*, 6059.
- [68] R. Tran, J. Lan, M. Shuaibi, B. M. Wood, S. Goyal, A. Das, J. Heras-Domingo, A. Kolluru, A. Rizvi, N. Shoghi, et al., *ACS Catal.* **2023**, *13*, 3066.

Manuscript received: January 25, 2024

Revised manuscript received: March 21, 2024

Accepted manuscript online: March 22, 2024

Version of record online: May 27, 2024