THE JOURNAL OF PHYSICAL CHEMISTRY C
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

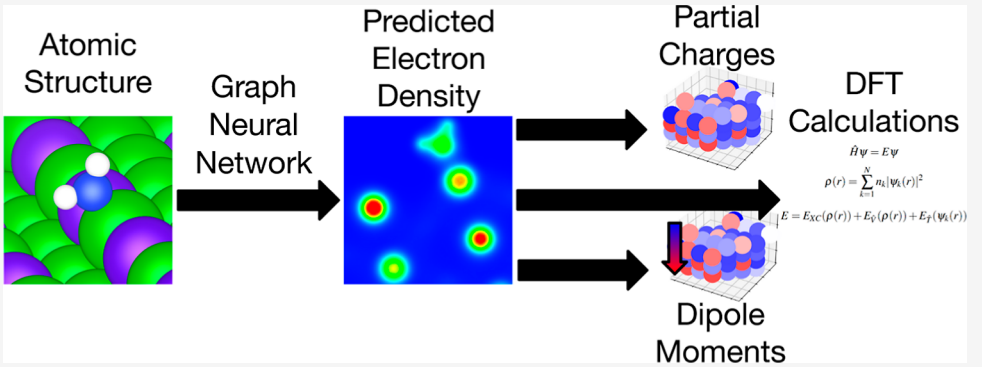
pubs.acs.org/JPCC
Article

Chemical Properties from Graph Neural Network-Predicted Electron Densities

Published as part of The Journal of Physical Chemistry C virtual special issue " Machine Learning in Physical Chemistry Volume 2".

Ethan M. Sunshine, Muhammed Shuaibi, Zachary W. Ulissi, and John R. Kitchin*

Cite This: J. Phys. Chem. C 2023, 127, 23459–23466



ABSTRACT: According to density functional theory, any chemical property can be inferred from the electron density, making it the most informative attribute of an atomic structure. In this work, we demonstrate the use of established physical methods to obtain important chemical properties from model-predicted electron densities. We introduce graph neural network architectural choices that provide physically relevant and useful electron density predictions. Despite not being trained to predict atomic charges, the model is able to predict atomic charges with an error of an order of magnitude lower than that of a sum of atomic charge densities. Similarly, the model predicts dipole moments with half the error of the sum of the atomic charge densities method. We demonstrate that larger data sets lead to more useful predictions for these tasks. These results pave the way for an alternative path in atomistic machine learning where data-driven approaches and existing physical methods are used in tandem to obtain a variety of chemical properties in an explainable and self-consistent manner.

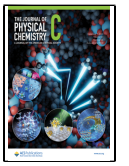
INTRODUCTION

Global implementation of sustainable energy technologies requires the discovery of efficient and economical catalysts for chemical reactions, such as water oxidation, carbon dioxide reduction, and ammonia synthesis. In these reactions, it is common to use a solid surface as a heterogeneous catalyst. Because the design space of material surfaces is extremely vast, computational modeling techniques are employed to identify promising catalysts for further experimental study. Broadly speaking, these methods can be divided into two categories: physics-based and data-driven.

First, there are physics-based methods. Of these, density functional theory (DFT) is the most popular choice for heterogeneous catalysis. In DFT, the valence electrons in a chemical structure are treated as a density field rather than individual particles described by wave functions. This method is based on the two Hohenberg–Kohn theorems, which state that (1) all chemical properties are a functional of electron density and (2) the true electron density minimizes the energy functional.[1,2] Thus, in a typical DFT calculation, the initial electron density is guessed. This density is updated iteratively until the energy is minimized. Many properties can be easily calculated from the converged electron density, such as atomic charges, dipole moments, and atomic forces.

More recently, machine learning methods have been developed as computationally cheaper alternatives to DFT. One popular machine learning method for chemistry is the graph neural network (GNN). In this paradigm, atoms are represented as nodes in a graph, and information is exchanged along the edges between them in a scheme called "message

Received: September 13, 2023
Revised: October 29, 2023
Accepted: November 9, 2023
Published: November 27, 2023



https://doi.org/10.1021/acs.jpcc.3c06157
J. Phys. Chem. C 2023, 127, 23459–23466

passing." Assuming a constant volume, the computational cost of GNNs scales with the square of the number of atoms, while the cost of DFT generally scales with the cube. When assuming a constant atomic density, the comparison is even more favorable for GNNs: the cost scales linearly with increasing volume for GNNs but cubically for DFT.

However, in comparison to DFT, GNNs are yet to demonstrate competitive accuracy in heterogeneous catalysis. The largest and most diverse data set for heterogeneous catalysis is Open Catalyst 2020 (OC20).[3] Each structure in this data set is an adsorbate placed at a site on a catalyst surface, and the target is the associated adsorption energy. However, as of yet, no machine learning model achieves an adsorption energy mean absolute error (MAE) below 0.3 eV (29 kJ/mol) in the test data.[4] By propagating this uncertainty through a simple Arrhenius model at a temperature of 300 K, one can see an uncertainty of more than 5 orders of magnitude in kinetic rate constants. Thus, there is a pressing need for more accurate models in this field.

To bridge the gap between the accurate physics-based models and machine learning models, machine learning can be used to predict electron density directly. An accurate prediction of the electron density can be used in tandem with established methods, including DFT, to predict many physical quantities. For example, the Hellmann−Feynman theorem asserts that electron density alone is sufficient to determine atomic forces without the need to assume any particular functional.[5,6] By integration of these physical methods, machine learning models can become more useful, explainable, and transferable.

Because the electron density exists in three dimensions throughout the atomic structure, a significant amount of information is required to represent it accurately. It can be projected onto a basis set, but more often, in heterogeneous catalysis, electron density data are available at a grid of points. The density is sampled at a large number of points (in our data, 5×10^6 on average) throughout the atomic structure. This huge amount of data is largely underutilized by machine learning methods to date. In fact, most machine learned potentials do not use electron density data at all.

Several machine learning methods exist for electron density prediction. In 2018, Grisafi et al. demonstrated that Gaussian process regression can be used to predict electron density in terms of atom-centered basis functions.[7−10] This low-dimensional representation of electron density is natural for molecules. On the other hand, it is not trivial to project electron density onto atom-centered basis functions in periodic systems such as surfaces, where periodic basis functions such as plane-waves are common.

More recently, Rackers et al. introduced an equivariant GNN to predict electron density.[11,12] They also demonstrate that increasing the order of the atom-centered basis function set is necessary for accurate short-range force prediction. Increasing basis set order rapidly increases the computational cost of a GNN, as well as the difficulty in fitting an appropriate atom-centered basis to a plane-wave-based electron density. Considering these factors, this method is not well-suited for heterogeneous catalysis at present.

On the other hand, pointwise methods do not rely on atom-centered basis function representations of the electron density. Instead, the model predicts electron density at predetermined points in space called probes. This allows electron density prediction at arbitrary resolutions, and the cost increases only linearly with the number of probes. This method is particularly attractive for periodic systems because the electron density does not have to be represented by atom-centered basis functions. Furthermore, popular DFT codes for heterogeneous catalyses, such as the Vienna Ab Initio Simulation Package (VASP),[13−15] readily provide point-wise samplings of electron density. In the method introduced by Jørgensen and Bhowmik,[16] the probes are added to the graph in the GNN as information receivers (but not senders). One may also think of these probe points as "virtual atoms", which have been used in atomistic GNNs at least since the introduction of SchNet.[17] Similarly, Achar et al.[18] introduced a pointwise approach to electron density predictions for solid materials, although this method is not a GNN. We decided to use a pointwise approach in this work due to its natural integration with plane-wave DFT codes.

METHODS

Modeling. Our model is based on that of Jørgensen and Bhowmik. We make two key enhancements to the model that enable the use of its predictions with physical methods such as DFT. First, many probes are sufficiently far from any atoms, such that they receive no information in the message-passing procedure. Our model assumes the electron density at these points to be zero. The electron density decays slowly as the distance to the atoms increases, so probe points outside the cutoff radius often still carry some electron density. Since the model has no information about these points, it is forced to predict the same density for all of them. After training, this results in a nonphysical uniform negative charge in the void space, which can have undesired consequences when these predictions are used in downstream applications such as charge partitioning and DFT calculations. Instead, our model constrains this density to be zero.

Second, our model enforces a charge balance. Since our systems are periodic, a prediction where the total charge is not balanced exactly is unsuitable for many physical models. Thus, our model takes as an input the total number of valence electrons in the atomic structure. We implement the following assumption: a sample of probes drawn from a uniform random distribution should have a proportional share of the valence electron density. That is, if the probes are randomly divided into 10 batches, each batch should contain one-tenth of the valence electron density. This constraint is strictly enforced by a normalization step inside the model: after a message passing, the predicted density values are normalized to have a sum of one. Then, they are multiplied by the total number of valence electrons in the structure and the fraction of probes in the sample. This assumption is necessary because treating all probes in a single batch is computationally infeasible. It also means that after predictions are made on all batches, the total electron density in the atomic structure is guaranteed to sum to the correct number of valence electrons exactly. We observed that small batches of probes resulted in worse accuracy and unstable training. This is likely because this assumption does not apply to small batches.

Outside of the model differences, several additional features differentiate this work from previous work. Our implementation is built as an extension of the Open Catalyst Project (OCP) modeling software. This enables the integration of existing message-passing models, including the reuse of models trained on other tasks. For example, model weights from a machine-learning potential can be used. We anticipate that

https://doi.org/10.1021/acs.jpcc.3c06157
J. Phys. Chem. C 2023, 127, 23459−23466

further integration with this software will allow us to leverage future developments in the OCP code repository.

Another notable feature is the probe edge-finding algorithm. Previous implementations had a computational cost of $O((A + P)^2)$, where A is the number of atoms and P is the number of probes. This is because no distinction was drawn between atom and probe nodes, and a general neighbor-finding algorithm was used. Because we are not interested in edges between probes, the cost of this algorithm can be reduced to $O(A(A + P))$, which is significantly better because A is several orders of magnitude smaller than P. To achieve this, we modified the neighbor-finding algorithm available in the OCP modeling software. The standard procedure is to consider every possible edge between two nodes, determine the distance, and remove any that are longer than some specified cutoff radius. Because we are not interested in probe−probe edges, we can save a significant amount of memory by not computing these distances. We partition the graph into two sets: the set of atoms and the set of probes. Distances are computed only from nodes in the set of atoms. This procedure is written in Pytorch code that runs efficiently on both the CPU and graphics processing unit (GPU). It also respects periodic boundary conditions in one, two, or three directions. Under previous implementations of the probe graph training procedure, the number of probes used in the training was limited by graph construction speed. In contrast, our implementation is limited by the memory requirements of a message passing through large graphs. This also enables rapid inference with a significant speedup compared to previous implementations.

Data Set. We prepared a large and diverse data set of charge densities for heterogeneous catalysis. The atomic structures in this data set are from OC20.[3] Each structure is composed of an adsorbate placed on a solid material surface. The training set consists of 32794 random structures from the OC20 train split, and the validation set consists of 897 random structures from the OC20 " both out of distribution" split. One hundred random structures from the OC20 "both out of distribution" split comprise the test data set. No adsorbate or surface that is present in the training set is present in the validation or test set. This data set includes 56 different chemical elements (Ag, Al, As, Au, B, Bi, C, Ca, Cd, Cl, Co, Cr, Cs, Cu, Fe, Ga, Ge, H, Hf, Hg, In, Ir, K, Mn, Mo, N, Na, Nb, Ni, O, Os, P, Pb, Pd, Pt, Rb, Re, Rh, Ru, S, Sb, Sc, Se, Si, Sn, Sr, Ta, Tc, Te, Ti, Tl, V, W, Y, Zn, and Zr), and has system sizes that range from 9 to 225 atoms. Specific details for each structure are included in a set of ASE database files in the code repository that is described in the Supporting Information. Self-consistent electronic structure optimization was performed in VASP using the Revised Perdew−Burke−Ernzerhof functional using the settings outlined for the construction of OC20.[3]

The electron density in this data set is obtained directly from the CHGCAR output file and represents a pseudovalence electron density. The density is sampled on a uniform, three-dimensional grid of points. The number of electron density probe points varies depending on the size of the unit cell. On average, each atomic structure in this data set contains about 5 million probe points.

Training Procedure. The models are trained to minimize the normalized MAE introduced by the authors, Jørgensen and Bhowmik. We note that there is no universally accepted error metric for the electron density. We acknowledge that this metric may not translate well across data sets and is biased toward high-density probe points. The metric is defined as

$$\varepsilon_{\mathrm{MAE}} = \frac{\int_{\vec{r}\in V} |\hat{\rho}(\vec{r}) - \rho(\vec{r})|}{\int_{\vec{r}\in V} |\rho(\vec{r})|} \tag{1}$$

where ρ is a ground-truth electron density obtained from a converged self-consistent DFT calculation, $\hat{\rho}$ is a model-predicted electron density, and V is the simulation volume. This integral is evaluated numerically by assuming that probe points represent equal-volume voxels, each of uniform density that matches the density at the probe point.

At each training step, one atomic structure is chosen from the training data set and converted to a geometric graph structure. Then, a subsample of probes is randomly chosen from a uniform distribution. The number of probes chosen at this step can be thought of as the "batch size" and is generally limited by the GPU memory capacity. For these models, the batch size used for training is 10^4 or more. The geometric graph is augmented with the probe nodes and their respective edges. Then, the model makes a prediction of the electron density at each probe point. The error metric is computed, and back-propagation is used to obtain gradients for each model parameter. The model parameters are adjusted according to the Adam algorithm.[19] This procedure is repeated until each structure in the training data set has been chosen once, which comprises one epoch. The model is trained until no further improvements are seen on the validation set, which typically takes about 20 epochs.

In this work, we will focus on one model, a SchNet-based model,[17] which is akin to Jørgensen and Bhowmik's "Invariant DeepDFT". Our implementation is built from the OCP implementation of SchNet. The main architectural differences between these models and DeepDFT are charge conservation and the long-range zero density constraint, as described in the modeling section of this work. The hyperparameters of each of the models are available in the Supporting Information of this work.

Interface with DFT Calculator. To make predictions of physical properties such as forces, it is advantageous to use an existing quantum mechanical calculator. For instance, one may wish to perform a self-consistent field (SCF) calculation by using model-predicted electron densities as an initial guess. Thus, we have established methods to use predictions from our electron density models as inputs to VASP. To do this, it is necessary to have electron density predictions on the same grid that is sampled in the CHGCAR file. Obtaining these predictions on a single GPU takes about 1−2 min for a typical structure in our data set.

The electron density prediction is written in a CHGCAR file. To use this file as a VASP input, one must also write projector-augmented-wave (PAW) occupancies. In this work, we reuse PAW occupancies obtained from a DFT calculation with zero electronic minimization steps. This non-SCF calculation is relatively fast and is necessary only to obtain a reasonable guess of the augmentation occupancies. In principle, one may obtain reasonable augmentation occupancies by another method, such as reading the pseudopotential information directly. Once a CHGCAR is written with the atomic structure, electron density, and PAW occupancies, VASP can read the file with the INCAR flag ICHARG = 1. This causes VASP to use the predicted density as an initial

https://doi.org/10.1021/acs.jpcc.3c06157
J. Phys. Chem. C 2023, 127, 23459−23466

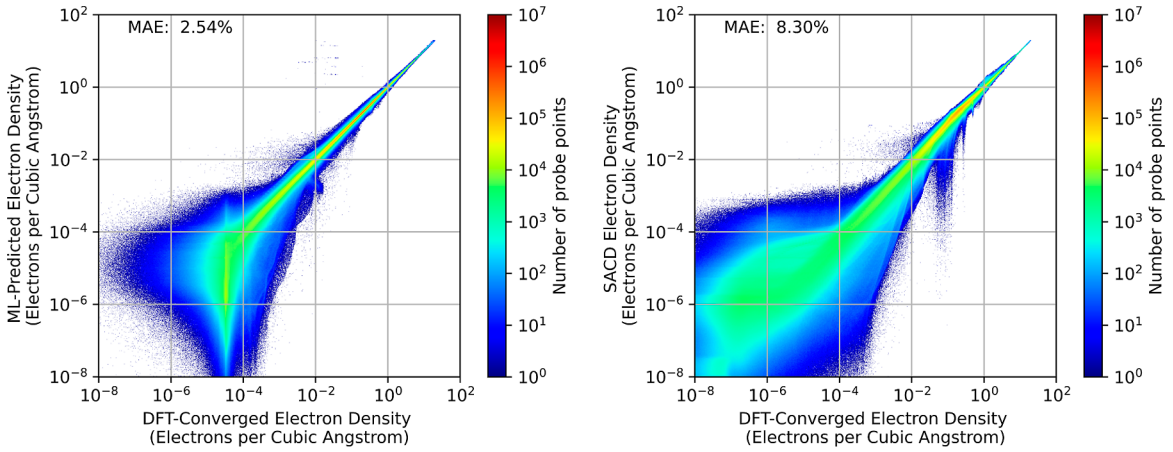


Figure 1. Parity heat maps comparing the results of the machine learning model and the SACD method. The model achieves tighter parity in the high-density region but not necessarily in the low-density region. This plot does not show points with zero or negative pseudovalence electron density.

guess for the electronic minimization loop instead of the default method.

The default method VASP uses to obtain an initial guess of electron density is a superposition of atomic charge densities (SACD). This is an easy and inexpensive way to obtain an electron density by adding the electron density that would be obtained from each atom if it were isolated. It ignores the effects of interatomic interactions on electronic structure. When discussing our results, we sometimes refer to the SACD method as a baseline.

RESULTS

Electron Density Predictions. Across the 100-structure test data set with 4.94 million probe points per structure, the model achieves an MAE of 2.54%. In comparison, the SACD method achieves an MAE of 8.30%. We show parity between the model-predicted and ground-truth electron densities on a parity heat map. This allows visualization of the model performance across different regimes in the test data set. Figure 1 shows the performance achieved by our best model. We see that our model achieves tighter parity in the high-density regime compared with the SACD method. This high-density regime includes many points in the data set, as evidenced by the thin warm-colored line in this portion of the plot.

Examining the low-density regime, it is difficult to claim that the machine learning model has significantly better or worse performance than the SACD method. It seems that the model is unable to effectively distinguish between points with electron densities below 10^{-4} electrons per cubic Angstrom. This is to be expected because such points have little impact on the error metric by which the model is trained. Furthermore, many points in this regime seem to have disappeared from the plot. This is because these points are beyond the cutoff radius of any atoms; therefore, the model is forced to predict zero electron density. Thus, these points are not shown in the logarithmic plot. Ultimately, precision in this low-density regime is likely unnecessary for many applications.

Learning Curve. To investigate the limitations of the model, we trained three additional models on smaller subsets of the training data set. The performance of each trained model is shown in Figure 2. From this learning curve, it appears that we are still in a regime where the accuracy of the model is limited not by the model architecture itself but by the size of the training data set.

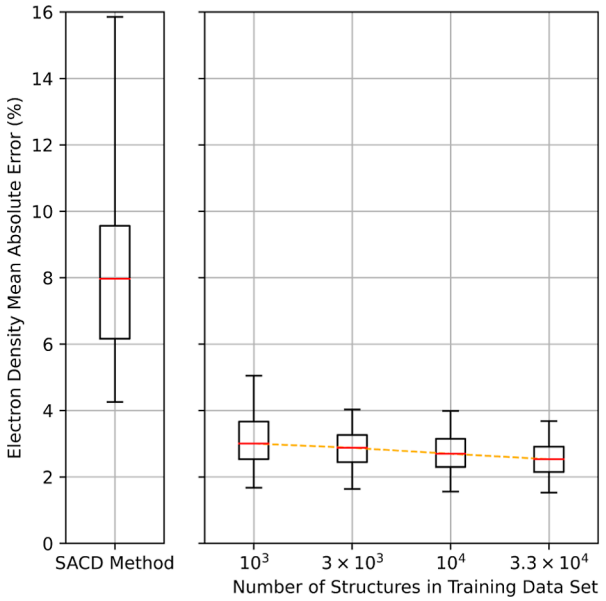


Figure 2. Learning curve with respect to the training set size. Boxes represent 50% of the test data, and whiskers represent 90%. Further increases in data set size are likely to result in even lower error measurements.

DFT Initialization. One potential use of these model-predicted electron densities is to initialize self-consistent DFT calculations. Because these calculations gradually converge an electron density guess to its ground state, it stands to reason that an initial guess that is close to the actual ground-state electron density will result in faster convergence. In reality, this is only one piece of the self-consistency puzzle. In a PAW calculation, it is necessary to converge the wave functions and augmentation occupancies in addition to the electron density. Thus, even a perfect guess of electron density will require a significant amount of electronic minimization steps to obtain self-consistency if the initial guesses of wave functions and augmentations are poor. Despite this limitation, we find that improving the error in electron density improves the overall

https://doi.org/10.1021/acs.jpcc.3c06157
J. Phys. Chem. C 2023, 127, 23459−23466

convergence speed of a DFT calculation. This result is demonstrated in Figure 3.

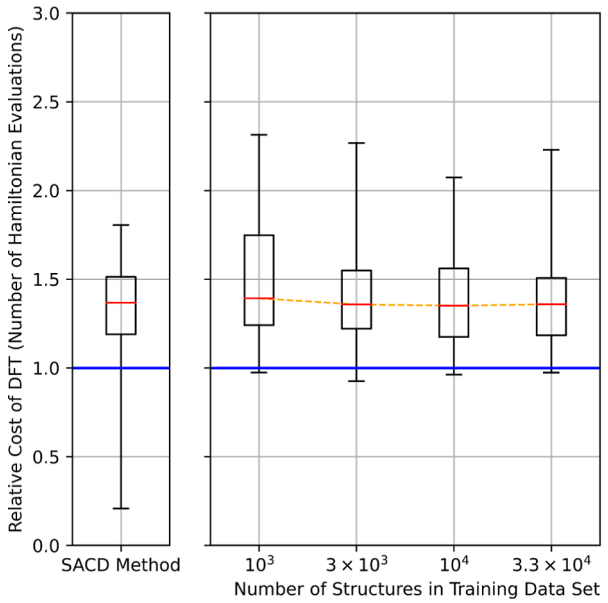


Figure 3. VASP performance is improved as the training set size increases. This is attributed to more accurate electron densities. The number of Hamiltonian evaluations is normalized by the number required when using the ground-state (i.e., perfectly accurate) electron density as an initial guess. Thus, the blue line represents the "best case scenario" for a perfect electron density model. Significant progress is still needed to approach this limit. Interestingly, the machine learning models do not consistently outperform the baseline SACD method in this task. Boxes represent 50% of the test data, and whiskers represent 95% of the test data.

This was measured by running a DFT self-consistency calculation (single point) on each structure in the test data set. This was repeated with the electron densities predicted by each model. Additionally, we ran the same experiment for the initialization of the actual ground-state electron densities obtained from DFT. To measure performance, we take the total number of evaluations of the Hamiltonian acting on a wave function, which is reported in VASP's output as "ncg." The performance is then normalized relative to the case in which the ground-state initialization is used as the initial guess.

The figure shows a clear correlation between the training set size and the resulting DFT performance. However, it is also clear that this scheme offers no practical value at this time. Even the best model presented in this work is unable to consistently yield electron density guesses that lead to faster convergence than that of the SACD method. It appears that simply increasing the training set size is not a practical path toward optimal performance on this task.

Atomic Charge Predictions. Another common use of electron density is to determine atomic charges. This is done through a charge partitioning scheme. These methods divide the structure into volumes, each of which contains one atom. Then, the total charge within that volume is summed, including the negative charge from the electron density and the positive charge from the nucleus. These total charges are assigned to each atom. Quantitatively, atomic charges can be used to investigate electrostatic interactions between atoms.[20] Qualitatively, they can help explain the movement of electrons between atoms and investigate the chemical reaction mechanisms. Furthermore, after dividing the space into these volumes, additional topological analysis can be performed to study chemical bonds.[21]

One common such scheme is Bader charge partitioning, in which volumes are defined by zero-flux surfaces.[22−24] We performed this analysis on the electron densities predicted by each of our models on the test data set, as well as the ground-state and SACD electron densities.[25] We evaluate the performance of each model by calculating the absolute error

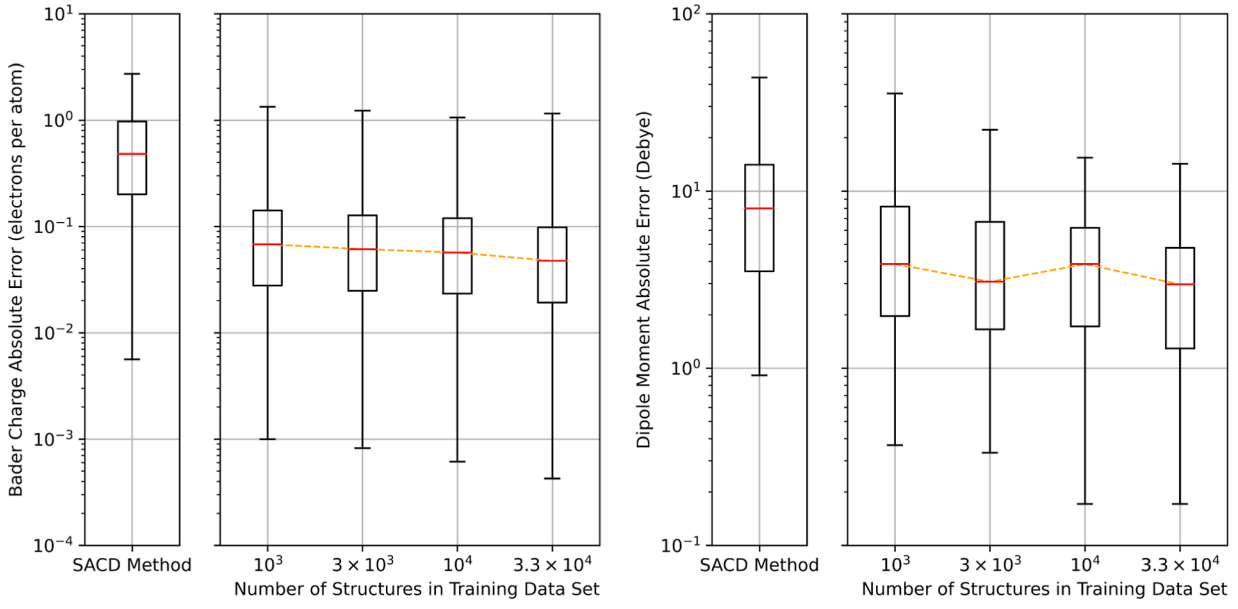


Figure 4. Error in Bader charges and electrostatic dipole moments. In each case, the machine learning models offer a marked improvement over the SACD method. Boxes represent 50% of the test data, and whiskers represent 95% of the test data. Training on larger data sets appears to offer slight improvements. Note that the models are not trained to predict either of these properties directly. Rather, existing methods are used to calculate them from the predicted electron densities.

https://doi.org/10.1021/acs.jpcc.3c06157
J. Phys. Chem. C 2023, 127, 23459−23466

between each predicted atomic charge in the test data set compared to the charges obtained from the ground-state electron density. The results are summarized in Figure 4.

The model performs significantly better than the SACD method, and the performance is improved with the size of the training data set. Several previous works have used message-passing GNNs for atomic charge prediction and achieved significantly better performance than this.[26−28] For instance, Metcalf et al. achieved errors near 10^{-2} electrons per atom via an electron passing scheme.[28] However, it is worth considering that, unlike previous methods, the method presented here does not include atomic charges in the training data. In this work, we are able to achieve this performance simply by training on electron density data and using existing methods to determine atomic charges. This enforces a conservation of the total charge, which is not guaranteed by all existing methods. Moreover, it demonstrates that accurate electron density predictions can be used to obtain other properties using established physics-based methods.

Dipole Moment Predictions. Electrostatic dipole moments were calculated from the electron densities and pseudo-ion charges. Like the Bader analysis, this calculation was done with the SACD density, ground-state density, and each model-predicted density for each structure in the test data set. The result is shown in Figure 4. The results show a trend similar to that of the Bader analysis. There is a marked improvement in the performance of the machine learning models compared to the SACD method, and training on larger data sets results in further decreases in error.

Assessing the Long-Range Zero Density Assumption. The assumption of zero density over long ranges is a possible source of error in the resulting predictions. The following experiments were conducted to evaluate the consequences of this assumption. For the test set, charge densities were constructed to be the ground-state density within 6 Å of an atom and zero density outside this region. This density represents the best possible model of electron density, given the long-range zero density assumption. We found that the MAE for this constructed density is 2.00%. The median dipole prediction error is 0.025 D. The median Bader charge error is 2×10^{-6} electrons per atom. This demonstrates that the long-range zero density assumption does not limit the model's ability to make useful predictions. It also shows that although the MAEs of the trained models are near this bound of 2.00%, there is significant room for improvement in the downstream chemical predictions.

DISCUSSION

Current Limitations. With any machine learning method, the resultant model is greatly limited by the training data that is used in its generation. As such, although the model can be used as-is to obtain predictions for out-of-domain systems, these predictions have no guaranteed degree of accuracy. Such out-of-domain systems include but are not limited to oxide materials, zeolite materials, metal−organic frameworks, amorphous materials, and liquid systems. It is possible to use the tools and techniques presented in this work to create an equivalent model for each of these domains, provided sufficient training data is available. Similarly, we have only trained on the pseudovalence electron density (available in the CHGCAR) and make no claims about model performance with all-electron density or similar quantities.

This model is not accurate enough to directly use the predicted electron density to obtain an estimate of forces, as has been done by Rackers et al.[11] We attribute this to the diversity of our data set in comparison to that used in other work, and we hope that future training on larger but equally diverse data sets will achieve sufficient accuracy for this task.

Predicting the full electron density of a structure in this data set with this model takes about 1−2 GPU-minutes. This is because the number of probe points is too large for the GPU to handle in one batch. Instead, the probe points are divided into many batches that are transferred to the GPU in succession. In comparison, machine-learned potentials can be evaluated in a fraction of a second. Bridging this speed gap will require optimizations, such as predicting fewer probe points and efficiently up-sampling to a full grid.[29]

Paths toward Improved Performance. As shown in Figure 2, it appears that model performance can still be improved by training on larger data sets. It is relatively easy to generate an electron density data set that is 1 or 2 orders of magnitude larger using DFT. Based on the learning curve, it seems that this would lead to models with an MAE near or below 2%. However, training on such a data set remains a challenge. Practically speaking, the current size of the data set is well over 500 GB. This is comparable in size to the full OC20 data set, despite containing 4 orders of magnitude fewer structures. In fact, due to the batching procedure, the models presented here have not truly been trained on every data point. An electron density data set with the same number of structures as OC20 would contain petabytes of data, placing it among the largest atomistic data sets created to date. Training on such large data sets requires extensive GPU resources and specialized parallelization techniques. Thus, there is a need for efficient implementations for storing and training on such large data sets. This path shows the most promise to improve the accuracy of these models.

On the other hand, one might consider training with a more advanced GNN model. The SchNet model used in this work is not considered state-of-the-art in OC20.[4,30] PaiNN-based architectures have already been used for electron density prediction.[16] This could also become very computationally expensive as more and more advanced architectures are used. Furthermore, not all GNN architectures are suitable for this task. For instance, a model such as GemNet that enumerates three- and four-body interactions[31,32] is highly impractical since a typical structure will have millions of "bodies": the probe points. It is critical to choose or create GNN message-passing architectures with a cost that scales linearly (or better) with the number of probe points in the batch.

Finally, we discuss some practical paths toward improved DFT performance. As mentioned previously, a self-consistent PAW calculation requires wave function and augmentation occupancies in addition to the electron density. Although it is possible to use the default methods for these quantities, improving the initial guesses of these properties will amplify any performance gains from an accurate electron density guess. Indeed, machine-learned wave functions have demonstrated DFT acceleration for nonperiodic systems.[33,34] We expect it should be possible to model the augmentation occupancies using equivariant neural networks that make atom-centered predictions (such as the electron density models introduced previously).[7−12] Predicting the periodic wave functions promises further SCF convergence speed improvements, but it is not immediately obvious how this can be efficiently

https://doi.org/10.1021/acs.jpcc.3c06157
J. Phys. Chem. C 2023, 127, 23459−23466

integrated with current GNN models. Further increases in the training set size are not likely to result in large performance improvements on this task. Further investigation is necessary to determine why the SACD method outperforms the machine learning models on this task despite significantly worse error metrics.

CONCLUSIONS

In this work, we implemented a message-passing GNN for the electron density prediction. It includes several useful features that enable the predicted electron density to be used with physical models, such as DFT and Bader charge partitioning. Although fast and accurate machine learning models exist for many individual properties, this work demonstrates an alternative approach in which models can be trained on electron density only and existing methods can be used to determine other quantities. We believe that training on larger data sets and with more advanced models will enable even more accurate models in this domain.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcc.3c06157.

More details about the code and data availability, as well as the hyperparameters of the models and training procedures presented in this work (PDF)

AUTHOR INFORMATION

Corresponding Author

John R. Kitchin − *Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States;* orcid.org/0000-0003-2625-9232; Email: jkitchin@andrew.cmu.edu

Authors

Ethan M. Sunshine − *Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States;* orcid.org/0000-0002-1636-9094

Muhammed Shuaibi − *Fundamental Artificial Intelligence Research, Meta Platforms, Inc, Menlo Park, California 94025, United States*

Zachary W. Ulissi − *Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States; Fundamental Artificial Intelligence Research, Meta Platforms, Inc, Menlo Park, California 94025, United States;* orcid.org/0000-0002-9401-4918

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcc.3c06157

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge support from Meta Platforms, Inc. via the Open Catalyst Project collaboration. We thank Arghya Bhowmik and Peter Bjørn Jørgensen for helpful discussions regarding the DeepDFT code. We are also grateful for discussions with Open Catalyst Project team members: Jehad Abed, Lowik Chanussot, Abhishek Das, Adeesh Kolluru, Janice Lan, Kyle Michel, Joseph Musielewicz, Saro Passaro, Ammar Rizvi, Nima Shoghi, Anuroop Sriram, Matt Uyttendaele, Brook Wander, Brandon Wood, and C. Lawrence Zitnick.

REFERENCES

(1) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864−B871.

(2) Sholl, D. S.; Steckel, J. A. *Density functional theory: a practical introduction*; Wiley: Hoboken, NJ, 2009.

(3) Chanussot, L.; Das, A.; Goyal, S.; Lavril, T.; Shuaibi, M.; Riviere, M.; Tran, K.; Heras-Domingo, J.; Ho, C.; Hu, W.; et al. Open Catalyst 2020 (OC20) Dataset and Community Challenges. *ACS Catal.* **2021**, *11*, 6059−6072.

(4) Lan, J.; Palizhati, A.; Shuaibi, M.; Wood, B. M.; Wander, B.; Das, A.; Uyttendaele, M.; Zitnick, C. L.; Ulissi, Z. W. AdsorbML: a leap in efficiency for adsorption energy calculations using generalizable machine learning potentials. *npj Comput. Mater.* **2023**, *9*, 172−179.

(5) Politzer, P.; Murray, J. S. The conceptual power of the Hellmann-Feynman theorem. *Struct. Chem.* **2023**, *34*, 17−21.

(6) Pathak, S.; López, I. E.; Lee, A. J.; Bricker, W. P.; Fernández, R. L.; Lehtola, S.; Rackers, J. A. Accurate Hellmann-Feynman forces from density functional calculations with augmented Gaussian basis sets. *J. Chem. Phys.* **2023**, *158*, 014104.

(7) Grisafi, A.; Fabrizio, A.; Meyer, B.; Wilkins, D. M.; Corminboeuf, C.; Ceriotti, M. Transferable Machine-Learning Model of the Electron Density. *ACS Cent. Sci.* **2019**, *5*, 57−64.

(8) Lewis, A. M.; Grisafi, A.; Ceriotti, M.; Rossi, M. Learning Electron Densities in the Condensed Phase. *J. Chem. Theory Comput.* **2021**, *17*, 7203−7214.

(9) Grisafi, A.; Lewis, A. M.; Rossi, M.; Ceriotti, M. Electronic-Structure Properties from Atom-Centered Predictions of the Electron Density. *J. Chem. Theory Comput.* **2023**, *19*, 4451−4460.

(10) Grisafi, A.; Bussy, A.; Vuilleumier, R. Predicting the Charge Density Response in Metal Electrodes. 2023, arXiv:2304.08966. http://arxiv.org/abs/2304.08966 [cond-mat] (accessed April 25, 2023).

(11) Rackers, J. A.; Tecot, L.; Geiger, M.; Smidt, T. E. A recipe for cracking the quantum scaling limit with machine learned electron densities. *Mach. learn.: Sci. Technol.* **2023**, *4*, 015027.

(12) Lee, A. J.; Rackers, J. A.; Bricker, W. P. Predicting accurate ab initio DNA electron densities with equivariant neural networks. *Biophys. J.* **2022**, *121*, 3883−3895.

(13) Kresse, G.; Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *47*, 558−561.

(14) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15−50.

(15) Kresse, G.; Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 11169−11186.

(16) Jørgensen, P. B.; Bhowmik, A. Equivariant graph neural networks for fast electron density estimation of molecules, liquids, and solids. *npj Comput. Mater.* **2022**, *8*, 183.

(17) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K. R. SchNet - A deep learning architecture for molecules and materials. *J. Chem. Phys.* **2018**, *148*, 241722.

(18) Achar, S. K.; Bernasconi, L.; Johnson, J. K. Machine Learning Electron Density Prediction Using Weighted Smooth Overlap of Atomic Positions. *Nanomaterials* **2023**, *13*, 1853.

(19) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. 2017, arXiv:1412.6980 [cs], http://arxiv.org/abs/1412.6980 (accessed July 2, 2023).

(20) Heinz, H.; Suter, U. W. Atomic Charges for Classical Simulations of Polar Systems. *J. Phys. Chem. B* **2004**, *108*, 18341−18352.

(21) Martiín Pendás, A.; Contreras-García, J. *Topological Approaches to the Chemical Bond; Theoretical Chemistry and Computational Modelling*; Springer International Publishing: Cham, 2023.

(22) Bader, R. F. W.; Henneker, W. H.; Cade, P. E. Molecular Charge Distributions and Chemical Binding. *J. Chem. Phys.* **1967**, *46*, 3341−3363.
(23) Tang, W.; Sanville, E.; Henkelman, G. A grid-based Bader analysis algorithm without lattice bias. *J. Phys.: Condens. Matter* **2009**, *21*, 084204.
(24) Yu, M.; Trinkle, D. R. Accurate and efficient algorithm for Bader charge integration. *J. Chem. Phys.* **2011**, *134*, 064111.
(25) Arnaldsson, A.; Tang, W.; Chill, S.; Chai, W.; Anselm, R.; Henkelman, G. *Bader Charge Analysis.* http://theory.cm.utexas.edu/henkelman/code/bader/ (accessed July 27, 2023).
(26) Raza, A.; Sturluson, A.; Simon, C. M.; Fern, X. Message Passing Neural Networks for Partial Charge Assignment to Metal-Organic Frameworks. *J. Phys. Chem. C* **2020**, *124*, 19070−19082.
(27) Thürlemann, M.; Böselt, L.; Riniker, S. Learning Atomic Multipoles: Prediction of the Electrostatic Potential with Equivariant Graph Neural Networks. *J. Chem. Theory Comput.* **2022**, *18*, 1701.
(28) Metcalf, D. P.; Jiang, A.; Spronk, S. A.; Cheney, D. L.; Sherrill, C. D. Electron-Passing Neural Networks for Atomic Charge Prediction in Systems with Arbitrary Molecular Charge. *J. Chem. Inf. Model.* **2021**, *61*, 115−122.
(29) Shen, J.-X. *mp-pyrho*, 2022; https://zenodo.org/record/7227448 (accessed: July 17, 2023).
(30) Liao, Y.-L.; Wood, B.; Das, A.; Smidt, T. EquiformerV2: Improved Equivariant Transformer for Scaling to Higher-Degree Representations. 2023, arXiv:2306.12059 [physics], http://arxiv.org/abs/2306.12059 (accessed July 27, 2023).
(31) Gasteiger, J.; Becker, F.; Günnemann, S. GemNet: Universal Directional Graph Neural Networks for Molecules. *35th Conference on Neural Information Processing Systems*, 2021 https://openreview.net/pdf?id=HS_sOaxS9K- (accessed Oct 29, 2023).
(32) Gasteiger, J.; Shuaibi, M.; Sriram, A.; Günnemann, S.; Ulissi, Z. W.; Zitnick, C. L.; Das, A. GemNet-OC: Developing Graph Neural Networks for Large and Diverse Molecular Simulation Datasets. 2022 https://openreview.net/forum?id=u8tvSxm4Bs (accessed Oct 29, 2023).
(33) Schütt, K. T.; Gastegger, M.; Tkatchenko, A.; Müller, K. R.; Maurer, R. J. Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions. *Nat. Commun.* **2019**, *10*, 5024.
(34) Unke, O. T.; Bogojeski, M.; Gastegger, M.; Geiger, M.; Smidt, T.; Müller, K. R. SE(3)-equivariant prediction of molecular wavefunctions and electronic densities. *35th Conference on Neural Information Processing Systems*, 2021, https://openreview.net/forum?id=auGY2UQfhSu (accessed Oct 27, 2023).

https://doi.org/10.1021/acs.jpcc.3c06157
J. Phys. Chem. C 2023, 127, 23459−23466