CHEMISTRY OF MATERIALS

pubs.acs.org/cm

Article

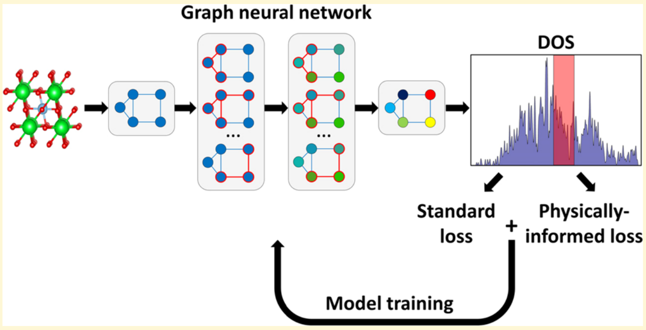
Physically Informed Machine Learning Prediction of Electronic Density of States

Victor Fung,* P. Ganesh, and Bobby G. Sumpter

Cite This: *Chem. Mater.* 2022, 34, 4848−4855

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: The electronic structure of a material, such as its density of states (DOS), provides key insights into its physical and functional properties and serves as a valuable source of high-quality features for many materials screening and discovery workflows. However, the computational cost of calculating the DOS, most commonly with density functional theory (DFT), becomes prohibitive for meeting high-fidelity or high-throughput requirements, necessitating a cheaper but sufficiently accurate surrogate. To fulfill this demand, we develop a general machine learning method based on graph neural networks for predicting the DOS purely from atomic positions, six orders of magnitude faster than DFT. This approach can effectively use large materials databases and be applied generally across the entire periodic table to materials classes of arbitrary compositional and structural diversity. We furthermore devise a highly adaptable scheme for physically informed learning which encourages the DOS prediction to favor physically reasonable solutions defined by any set of desired constraints. This functionality provides a means for ensuring that the predicted DOS is reliable enough to be used as an input to downstream materials screening workflows to predict more complex functional properties, which rely on accurate physical features.

■ INTRODUCTION

Electronic structure calculations are ubiquitous in modeling materials at the atomic scale and related to a broad range of fundamental and functional materials properties.[1,2] Density functional theory (DFT) is the most prevalent of these electronic structure methods, having been used to populate the majority of existing computational materials databases such as the Materials Project, OQMD, AFLOW, JARVIS, and others.[3−6] With DFT, the electronic density of states (DOS) can be obtained from the ground state electron density, which not only is a key intrinsic property of a material but also underlies many of its functional properties when used as electronic and optical devices, sensors, catalysts, and for energy storage, among others.[7−11] Therefore, the DOS, and its derived quantities such as band gaps and edge positions are heavily utilized in screening for promising candidates in many high-throughput materials discovery workflows, for example, in searching for suitable photoanodes for solar fuel generation.[12,13] In a recent study, the DOS was also used as an electronic fingerprint of materials undergoing structural transitions to understand the underlying causes for changes in electrical conductivity.[14] Additionally, the DOS can serve as an information-rich input to machine learning (ML)-based approaches for predicting more complex properties. In one such study, band edges were used as descriptors for an ML model to predict the impurity properties in semiconductors.[15] For catalysis, DOS descriptors such as the d-band positions and moments are extensively used to accurately predict surface adsorption energies, and by extension, the catalytic performance of various materials systems.[16−21] Extending this approach further, it was shown that the entire DOS range can be used as the input for a convolutional neural network model to accurately predict surface chemistry.[22] A shared characteristic in these applications is the need for fast and accurate evaluations of the DOS, something which is not always computationally tractable with DFT. Additionally, it is highly desirable for a method to provide the entire DOS rather than a specific derived value such as d-band center for catalysts, or band gap for optical devices, for maximum transferability and information retention. It has also been suggested that features derived from the full DOS can be more accurate than the same quantities predicted directly from the material descriptors.[23]

In this paper, we turn to ML to determine the DOS of a material from only the atomic positions, at a fraction of the cost of DFT, thereby addressing the major challenge of

Received: December 10, 2021
Revised: May 11, 2022
Published: May 24, 2022

https://doi.org/10.1021/acs.chemmater.1c04252

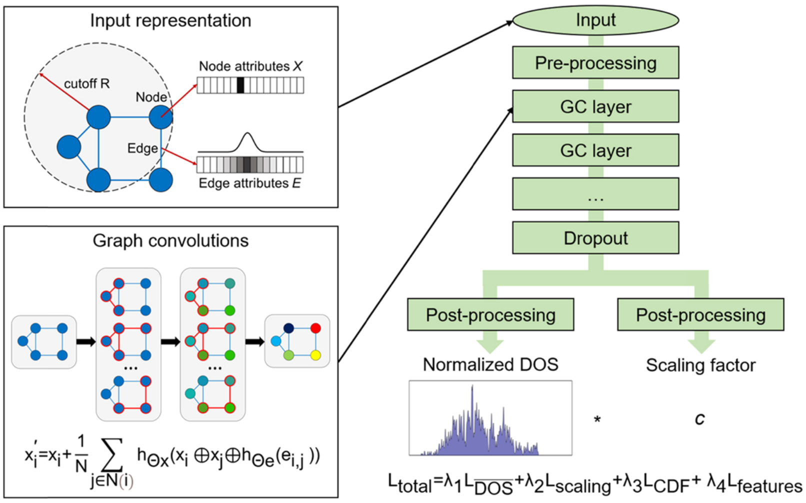


Figure 1. Overall schematic of the graph neural network model showing its individual components. (a) Input representation containing a graph with element one-hot node features and Gaussian-expanded interatomic distance edge features. (b) Graph convolutions which updates node embeddings from node and edge features from nearby neighbors. (c) Architecture of the graph neural network from input to the two output heads, and the loss function used to optimize the model.

efficient and accurate computation of the DOS. To accurately map atomic structure to the DOS with ML requires a representation which is sensitive to the full compositional and structural dimensions of the system. Furthermore, to obtain the projected DOS for each atom in the system, these representations should also be local in nature in order to distinguish changes in atomic environments. Structural representations such as atom-centered symmetry functions[24] and smooth overlap of atomic positions (SOAP)[25] can capture these changes in atomic environments for predicting the full atom-projected DOS, though these approaches do not incorporate the compositional dimension and remain limited to single element systems such as carbon or silicon.[14,23,26] Instead, we use a graph-based representation of the system containing atoms as nodes and interatomic distances as edges as inputs to a graph neural network[27,28] model to effectively capture both compositional and structural dimensions. Despite successful previous examples of graph neural networks in predicting materials properties,[29−32] the DOS is a high-dimensional target output which presents additional difficulties to be addressed such as smoothing, sensitivity to peak positions, and others.[23,26]

An important consideration that has not been addressed in current ML studies is a need to ensure that the predicted DOS also remains physically meaningful and accurate with respect to its derived features, such as band gaps and moments. In other words, given multiple predictions with an equal error in the DOS curve, the solution with the lower error in the derived features should be more preferred. This is particularly crucial when the predicted DOS is then used as an input to a subsequent, downstream, ML model to predict additional materials properties, where such a model would be highly sensitive to these features. Therefore, a method is needed to physically inform[33] the DOS prediction model to favor the more appropriate solutions. To tackle this challenge, we propose a general physics-informed approach for predicting the DOS which enforces the constraint within the model training process. A key advantage of this approach lies in its ability to use any set of differentiable DOS features to inform the training with minimal modifications to the model, as opposed to ad hoc methods hard coded into the model architecture. As different set features or descriptors are often required for the various application domains, whether it be catalysis or photovoltaics, this approach provides the greatest applicability across the domains.

■ METHODS

We use a graph neural network model which takes an input graph and returns a new graph containing updated node embeddings with information from its neighbors (Figure 1). The input graph, representing the atomic structure, uses a one-hot representation of atomic numbers for the node features and interatomic distances expanded by a gaussian basis for the edge features. Input node features first pass through a preprocessing step in the form of a dense layer, followed by graph convolution layers with convolutions taking the following form:

$$x_i' = x_i + \frac{1}{N}\sum_{j\in N(i)} h_{\Theta x}(x_i \oplus x_j \oplus h_{\Theta e}(e_{i,j})) \tag{1}$$

where x_i is the ith node in the graph, x_j is the neighbor, and $e_{i,j}$ is the edge connecting the two nodes. Here, $h_{\Theta x}$ and $h_{\Theta e}$ are single dense layers with PReLU activation functions. Batch normalization is then applied after each graph convolution layer. A dropout layer follows the last graph convolution layer. The resulting latent node embeddings then pass through two separate postprocessing layers, which provides an output each, the normalized DOS, and the scaling factor. The normalized DOS is a vector representing the discretized DOS, which has a length of 400 in this work, and the scaling factor is a scalar quantity.

The models are trained on the training set with the AdamW optimizer for 800 epochs, with the best model image chosen from the validation set. The performance is then evaluated with the test set. We use a data set split of 80% training, 5% validation, and 15% test. The dimension of the preprocessing and postprocessing layers, the number of graph convolution layers, the dropout ratio, the learning rate, and batch size are chosen through hyperparameter optimization with 200 trials for each data set. The best-performing hyperparameters are listed in the Supporting Information. The model construction and training was performed through a modified version of the MatDeepLearn code[31] which uses the pytorch and pytorch-geometric libraries.[34] The training time required for the $SrTiO_3$ (~5000 entries), alloy surface (~2000 entries), and bulk crystal data sets (~50,000 entries) on four V100 GPUs were 5, 4, and 90 min, respectively.

https://doi.org/10.1021/acs.chemmater.1c04252

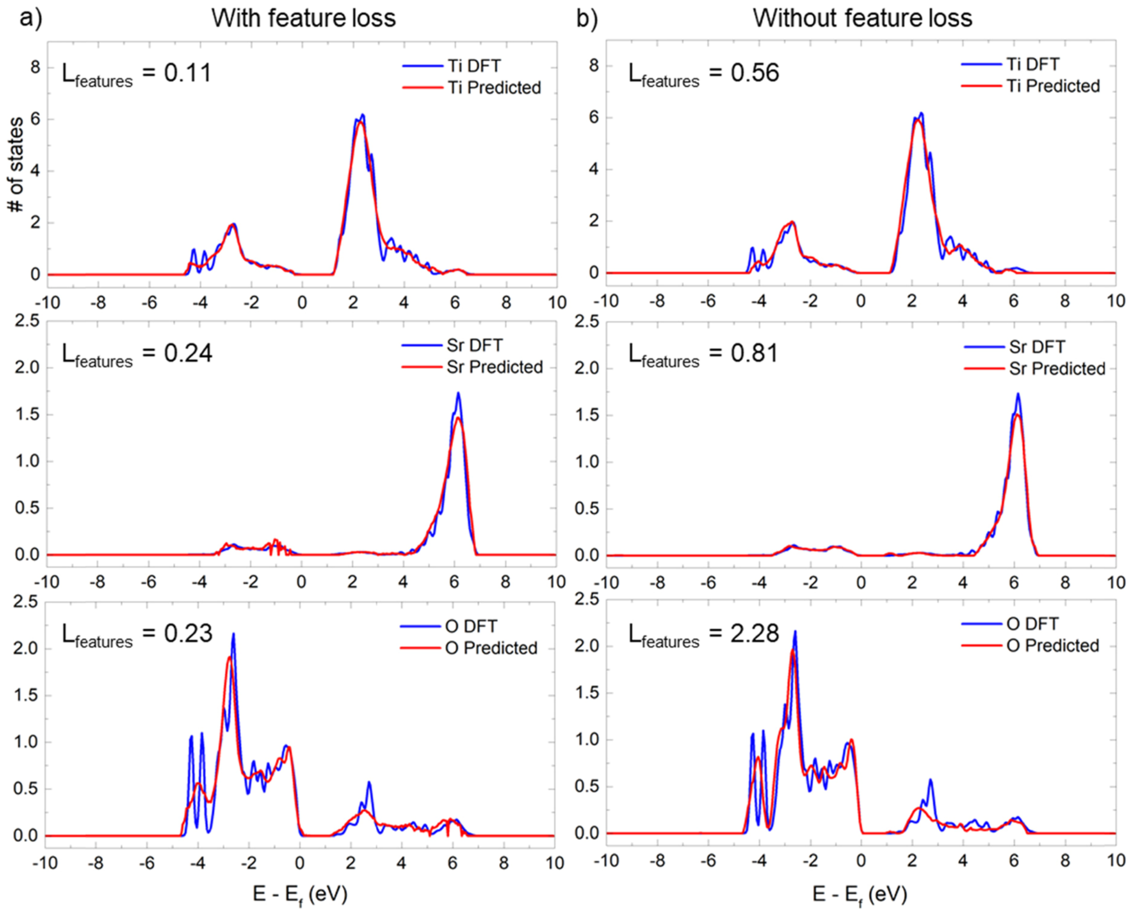


Figure 2. Density of states curves for a selected sample with median loss from the $SrTiO_3$ data set, showing predicted curves (red) overlaid on the DFT curves (blue). A side-by-side comparison is provided for a model trained on the same data set split (a) with L_{features} included and (b) without L_{features}.

In our model, we split the output containing the predicted DOS into two components, the min-max normalized DOS and the scaling factor c (Figure 1c). This design choice was made to prevent occurrences where atoms with a larger number of states end up dominating the training of the model, as their contribution to the loss is greater when the DOS is not normalized. This occurs when the atom-projected DOS is normalized with respect to the local charges per atom, leading to atoms with relatively few states. Without normalization, this resulted in predictions resulting in empty DOS for atoms with fewer total number of states. Consequently, we use a loss function which minimizes both the loss of the normalized DOS $L_{\overline{\text{DOS}}}$ as well as the scaling factor L_{scaling}, which is needed to reconstruct the original DOS. In addition, we include a third loss function, the loss of the cumulative distribution function of the normalized DOS L_{CDF}, utilized in a recent work in DOS prediction.[23]

Finally, in order to enforce the accuracy of not just the overall spectra but also its physical features, we include a fourth loss function L_{features}, which is the difference between the ground truth and the predicted features. This represents a general approach where the choice of the specific physical features can be chosen at will, usually depending on the desired target application. In other words, a learning bias is applied which penalizes unphysical features in the DOS.[33] As a proof of concept, in our current work, we choose as features the first four moments of the DOS, the band position, width, skew and kurtosis, and a fifth term, the number of states near the Fermi level. The quantities are computed from the DOS of the training data and serve as additional prediction targets. These features were chosen due to their importance in controlling the nature of surface-adsorbate interactions in surface chemistry and catalysis, and which are widely used as descriptors of adsorption energy.[16−21] Here, L_{features} is then the sum of the individual loss of the five features. With automatic differentiation, the gradients for the feature loss can be obtained and used to optimize the model in the same manner as the previous loss functions. The relative importance of the four loss functions is controlled by weights λ_1, λ_2, λ_3, λ_4 which can be tuned accordingly (in our current work, the weights are 1, 0.05, 0.005, and 0.15, respectively). Controlling these weights thereby offers a great degree of flexibility in controlling how strongly the physical constraints are to be enforced. For all loss functions, we used the mean absolute error (MAE) loss.

In addition, we tested two non-GNN models, using the SOAP[25] and Local Many Body Tensor[35] (LMBTR) descriptors for comparison. These methods provide local descriptors for each individual atom in the atomic structure, which are then fed into conventional fully connected neural networks to make predictions of the atom-projected DOS. The DScribe library was used to compute the SOAP and LMBTR descriptors.[36] All other aspects of these models follow the graph neural network models, such as training, loss functions, data splitting, and hyperparameter optimization. Finally, an overall baseline is provided in the form of a dummy regressor model which predicts a constant zero value for the DOS, to serve as a comparison between data sets.

We trained our models on three computational data sets, referred to as the $SrTiO_3$ data set, the alloy surface data set, and the bulk crystal data set. The $SrTiO_3$ data set contains 5000 bulk perovskite $SrTiO_3$ crystal structures which have undergone perturbations in its six lattice parameters a, b, c, α, β, γ. The parameters were sampled uniformly within a range of a 10% deviation from the equilibrium crystal parameters: $a = b = c = 3.914$ Å and $\alpha = \beta = \gamma = 90°$. The alloy surface data set contains 1913 bimetallic surface structures derived from a previous study compiling surface adsorption energies for catalyst screening.[37] All surfaces take the form of the geometry-

Table 1. Testing Loss of the DOS and Specified Features for the Investigated Data Sets[a]

	normalized DOS	band center	band width	band skew	band kurtosis	states near E_f
$SrTiO_3$ data set	0.046 (0.042)	0.052 (0.073)	0.154 (0.288)	0.024 (0.041)	0.056 (0.103)	0.022 (0.027)
alloy surface data set	0.102 (0.090)	0.094 (0.162)	0.294 (0.725)	0.075 (0.369)	0.188 (5.593)	0.133 (0.181)
bulk crystal data set	0.100 (0.098)	0.287 (0.330)	1.488 (1.888)	0.123 (0.187)	0.578 (1.270)	0.150 (0.161)

[a] Values in parentheses denote training without incorporating physical constraints, i.e., the features loss.

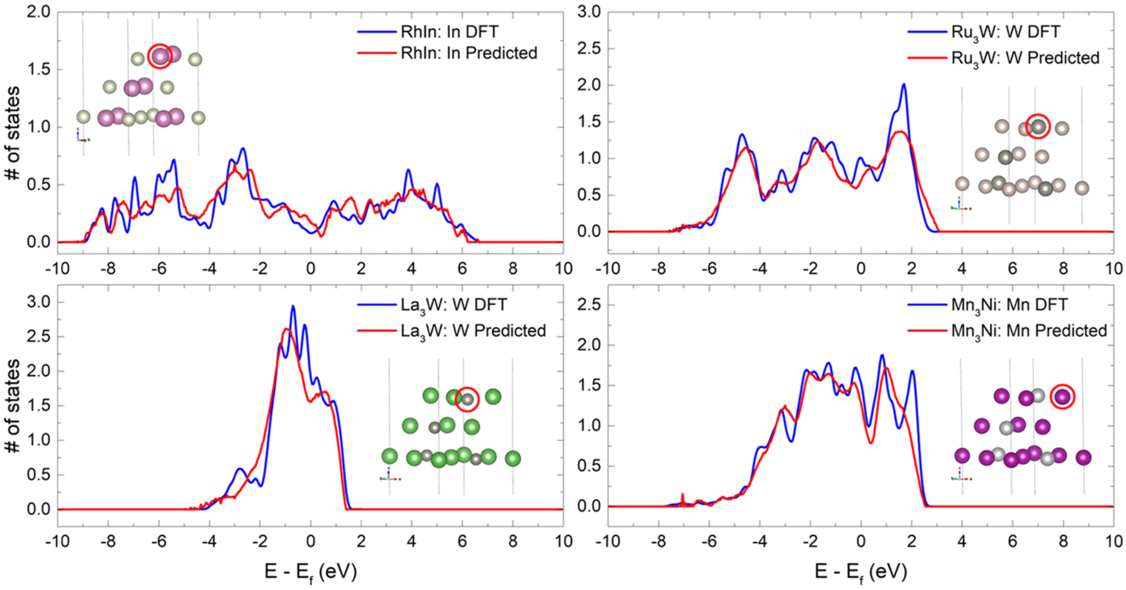


Figure 3. Representative density of states curves predicted with the graph neural network for the test set of the alloy surface data set with mean absolute errors at the 50th quantile, showing predicted curves (red) overlaid on the DFT curves (blue).

optimized (111) termination of the bulk fcc structure, but differ in their composition. The stoichiometry of the surfaces is either A, AB, or A_3B, where A and B are one of 37 sampled metallic elements in the periodic table. The bulk crystal data set is sourced from the Materials Project[6] and contains 50,789 bulk structures of inorganic crystals comprising a total of 88 elements. These structures are a subset of the full Materials Project database, and structures with a retrievable DOS with a sampling gridpoint of 2000, omitting ones with fewer gridpoints of 300 and 600.

We calculated the DOS of the $SrTiO_3$ data set in this work using DFT. The calculations were performed with the Vienna Ab Initio Simulation Package (VASP).[38,39] The Perdew−Burke−Ernzerhof[40] functional within the generalized-gradient approximation was used for electron exchange and correlation energies. The projector-augmented wave method was used to describe the electron−core interaction.[38,41] An on-site Coulomb interaction was included using the DFT + U method by Dudarev et al.[42] in VASP using a Hubbard parameter U = 4 eV for the Ti. A kinetic energy cutoff of 500 eV was used. All calculations were performed with spin polarization. The Brillouin zone was sampled using a Monkhorst-Pack scheme with a 8 × 8 × 8 grid.[43] The atom-projected DOS is then obtained for each structure. The DFT calculation details for the alloy surface[22] and bulk crystal[44] data sets can be found in their respective references. For three data sets, the raw DOS is preprocessed by smoothing with a gaussian filter, shifting the energy range from −10 to 10 eV with respect to the Fermi level and interpolating the curve to a resolution of 0.05 eV. This is done to ensure that the DOS curve has a consistent energy range and resolution within an external data set; however, this step can also be omitted if such a criterion is already enforced at the DFT calculation stage.

■ RESULTS

We first evaluated the performance of our method for DOS prediction on the $SrTiO_3$ data set. $SrTiO_3$ is a perovskite metal oxide with a computed band gap of 2.4 eV which undergoes an insulator to metal transition under applied strain or pressure. With three elemental components and a complex electronic behavior, this represents a challenging prediction task which furthermore requires a high degree of sensitivity with respect to atomic positions, as minor perturbations can induce significant changes in the band gap and overall DOS. Training on the data set with our model yielded a L_{DOS} of 0.046 states/eV over the entire test set. To better visualize the predictive performance, we plot an example of a predicted DOS from the test set in Figure 2a. We find that the DFT and predicted DOS curves match very closely with respect to overall distribution, and major peaks are reliably reproduced, though small peaks tend to be smoothed out. Visualizing the predicted DOS at various quantiles from 5th to 95th (Figures S1−S4) shows closely matched distributions in even the high error cases, and no obvious outliers were observed for this data set. A histogram of the errors in the test set is shown in Figure S29.

To investigate whether the inclusion of the features loss, L_{features}, will degrade the overall predictive performance, we plot the DOS curve of the same test sample trained on the same training split, but with the L_{features} set to zero (Figure 2b). We find the two cases to be visually similar beyond some minor variations, but the error in the DOS with respect to the features (L_{features}) is much higher when this loss term is not included. This is similarly reflected in the performance across the entire test set, where the accuracy with respect to the features is significantly worse (Table 1, values in parentheses). Overall, including L_{features} with a weight of 0.15 slightly increases the test set L_{DOS} by 9.5% while reducing the L_{features} by an average

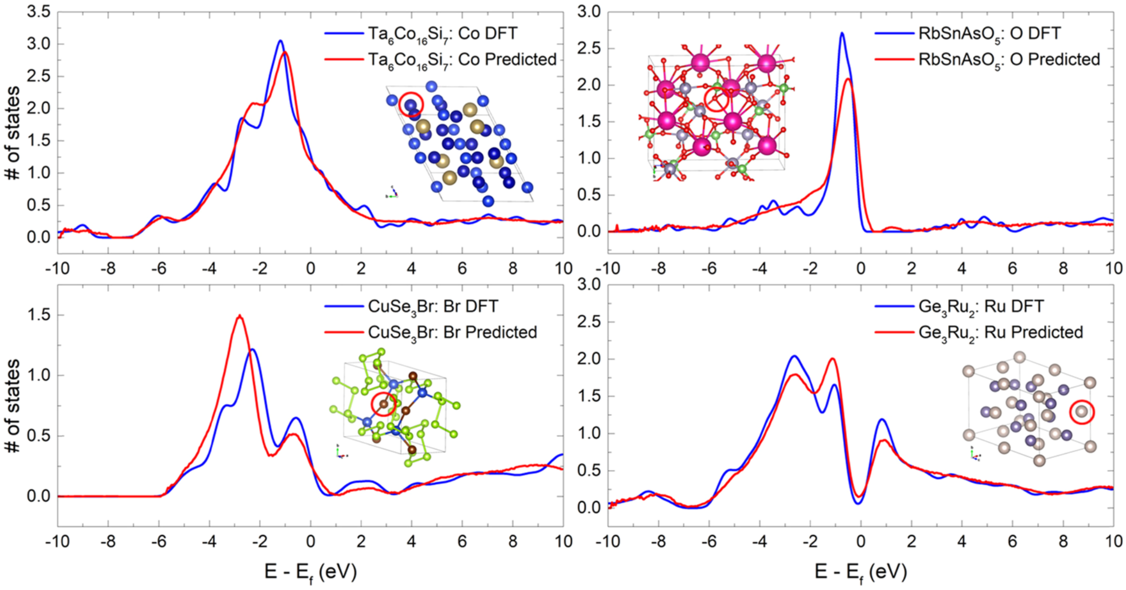


Figure 4. Representative density of states curves predicted with the graph neural network for the test set of the bulk crystal data set with mean absolute errors at the 50th quantile, showing predicted curves (red) overlaid on the DFT curves (blue).

of 36.2% for this data set, which can be considered an acceptable tradeoff.

Next, we evaluated the method for the alloy surface data set containing bimetallic alloy surfaces. This data set was featured in several recent studies where information from the DOS was used to predict adsorption energies for a given surface active site.[22,45] Here we obtained a higher test L_{DOS} of 0.102 states/eV, noting that this task is harder than the previous one due to having less than half the training data while going from three possible elements to 37. A histogram of the errors is shown in Figure S30. Despite the added complexity, the performance remains strong with the distribution of the predicted DOS matching closely with the DFT in Figures 3 and S13−S16. We can observe that the predictions remain accurate for different elements as well as for cases containing the same element but different neighboring environments, such as W in La_3W and Ru_3W (Figure 3). At higher error quantiles (Figures S15−S16), the height of the predicted DOS peaks tends to deviate, particularly for atoms with very narrow bands, but the number and position of these peaks remains well-matched. These results are promising as they suggest that this method works well even for relatively small data sets and with broad compositional ranges. Including L_{features} for this data set increases test L_{DOS} by approximately 13.3%, while reducing test L_{features} by an average of 60.8% (Table 1). The improvement is particularly dramatic for some of the higher order moments such as skew and kurtosis, where the occurrence of outliers and the overall error is reduced significantly.

To validate the effectiveness of this method in making general predictions of inorganic crystals of arbitrary structure and composition, we then evaluated the model on the bulk crystal data set, which is derived from the Materials Project database and contains 88 unique elements in total. This data set provides a reasonable sampling of the possible inorganic crystals encountered in nature and serves as a suitable benchmark for the goal of a general DOS prediction method for materials. Despite the significantly increased difficulty compared to the previous two examples, an L_{DOS} of 0.100 states/eV was achieved with our model. A histogram of the errors is shown in Figure S30. Plotting the median loss examples in Figure 4 again shows very closely matched curve distributions with only minor deviations in peak heights and locations. The performance remains strong up to the 75th error quantile (Figure S27) but degrades around the 95th quantile (Figure S28), with more noticeable changes in peak position and heights. This was anticipated due to the sparseness of the data set and the possibility that many atomic environments are not well-sampled. Nonetheless, extreme outliers were not observed, and predictions could still approximately match the overall distribution in the DOS. With regards to including L_{features}, the L_{DOS} increased by only 2.0% while reducing L_{features} by an average of 26.0% (Table 1). Hence, we find in this case that the tradeoff in curve prediction performance is almost negligible while providing a significant boost to the accuracy of its physical features.

Finally, to compare the performance of graph neural networks in this work with existing baselines using atomic-environment-based descriptors, we implemented and tested two additional neural networks which use SOAP[25] and LMBTR[35] atomic descriptors as inputs. A dummy model is also included in the benchmark which only predicts zero values for DOS to show relative differences between data sets. Evaluating the performance of these models on the three data sets in this work, we find that SOAP and LMBTR have slightly higher errors for $SrTiO_3$ but much higher errors for the alloy surface data set (Table 2). Predicted DOS curves for $SrTiO_3$ (Figures S5−S12) and surface alloy data sets (Figures S17−S24) are also shown, illustrating the degree of the degradation in predictive capability compared to the graph neural network model for the surface data set. This is largely unsurprising, as atomic-environment descriptors perform well for capturing spatial or geometric information, but becomes inefficient when the number of atom types becomes large.[46] For the bulk crystal data set containing 88 atom types, both SOAP and LMBTR representations become unwieldy and require hundreds of thousands of dimensions per atom. For this reason, the bulk

Table 2. Comparison of Testing Loss of the DOS Using Graph Neural Networks with Other Baselines

model	$SrTiO_3$ MAE	$SrTiO_3$ RMSE	alloy surface MAE	alloy surface RMSE	bulk crystal MAE	bulk crystal RMSE
this work	**0.046**	**0.095**	**0.102**	**0.258**	**0.102**	**0.201**
SOAP-NN	0.064	0.127	0.154	0.372		
LMBTR-NN	0.062	0.122	0.146	0.338		
dummy model	0.234	0.427	0.377	0.825	0.286	0.526

data set could not be evaluated due to memory limitations and extremely large model sizes.

We also investigate the data dependence of our model by plotting the training curves for each of the data sets in Figure 5, where the data split is kept constant and the proportion of training data is reduced from 80 to 60, 40, 20, and 10%. As we noted earlier, the alloy surface and bulk crystal data sets are relatively small given the compositional diversity of the materials, at approximately 2000 and 50,000 samples, respectively. For a comparison, ML data sets in molecular systems are generally much larger, with 130,000 samples for QM9,[47] while encompassing a much smaller elemental space (in this case containing only four elements, C, N, O, and F). As a result, we expect a significant room for improvement if the solid data sets are expanded to match the scale of molecular data sets. The training curves in Figure 5 appear to agree with this assertion, showing a continued decrease in the loss which will likely continue beyond the current data set sizes. Both the DOS and the feature losses exhibit similar trends of increasing performance with our approach. Extrapolating to one order of magnitude, more training data with a power law fit suggests that losses can still be reduced appreciably (Figure S32).

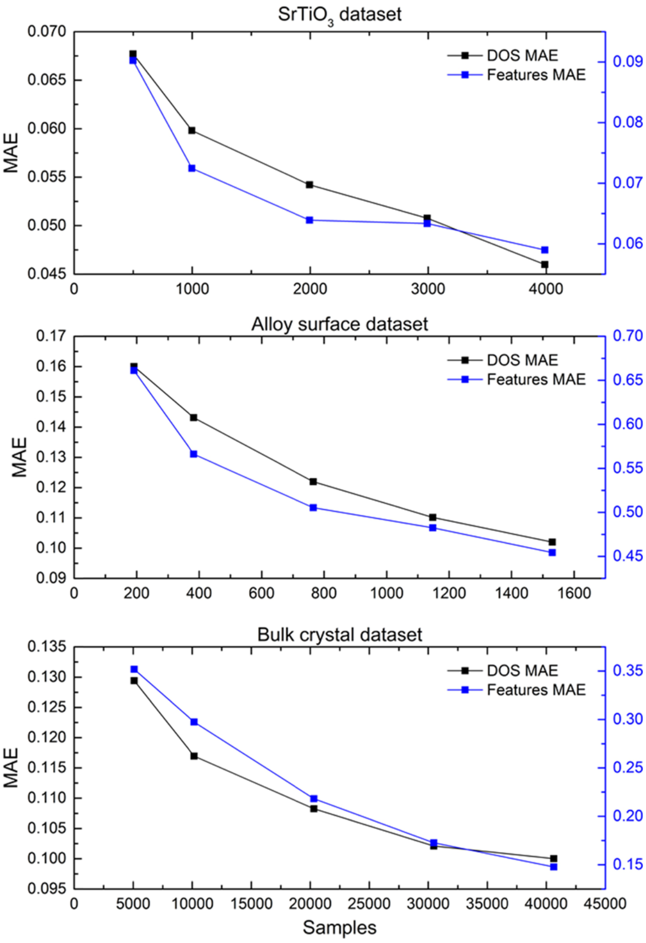


Figure 5. Training curves for models trained on the three data sets, where five points were obtained by using 10, 20, 40, 60, and 80% of the full data set for training, with the same split for validation and testing. Both the loss for the normalized DOS (black) and the features (blue) are shown here.

DISCUSSION AND CONCLUSIONS

The presented results provide a first glimpse into the accuracy of DOS prediction using a general ML framework with flexible physical constraints. This approach is able to effectively utilize large data sets and provides predictions at significantly faster speeds than DFT. For example, our model can predict the DOS of the full Materials Project database of 140,000 structures in 49 s on a single V100 GPU, or an average of 0.00036 s per crystal. This provides an unprecedented roughly six orders of magnitudes speedup over DFT, under the assumption that a single low-fidelity DFT calculation of the DOS will take several minutes per crystal, based on the timings from calculations performed in this work. This approach also scales linearly with system size, which allows for the DOS prediction of extremely large crystals or amorphous systems. With regards to model design, we note that there have been significant recent advances in graph neural network designs for materials chemistry applications.[48−50] These approaches can be incorporated in our framework with only minor changes to our overall workflow. Additional avenues for improvement can also include better output representations for the DOS, such as using a principal component basis[23,51,52] or using autoencoders,[53,54] which would help by reducing the output dimensions, as long as the reconstruction errors are sufficiently low. Alternatively, coupling the ML predictions to a physical model such as a tight-binding model or a lower level of DFT theory in a Δ-ML approach[55] could be used to improve accuracy and reduce data requirements.

We also presented a novel method for applying physical constraints through an additional loss function, which is used in training of the model via automatic differentiation to obtain gradients. An advantage of this approach is the flexibility of the desired constraints, which can be changed with no modifications to the rest of the workflow. For example, it is trivial to include other constraints such as band gap, states at the Fermi level, band energy, and the distribution of excitations,[23] as they are also computed from the DOS curve and are differentiable. While this may result in a tradeoff with the accuracy of the DOS curve, in many cases, the actual tradeoff is minor (~2% reduction in DOS accuracy for the bulk crystal data set), while retaining the full expressiveness of the original model. This generally offers advantages over physically informed ML schemes which enforces constraints by tailoring the ML architecture itself, which is often neither trivial nor generalizable and comes at the cost of model expressivity. Meanwhile, in other situations, where training data are limited, constrained model architectures which limit the functional space to physically meaningful solutions can prove to be beneficial.[56]

We anticipate the DOS prediction method to be especially useful as a source of computationally inexpensive electronic features for materials screening and ML studies which use the

DOS as the input to predict more complex functional properties. This removes the cost of the DFT from these studies, making it on par with methods which only use atomic positions as the input with respect to computational cost. The fast speed of the model also makes it suitable in providing electronic structure information during molecular dynamics simulations or for use as a method of informing experiments in real time in the form of a digital twin. As computational data sets containing DOS data continue to expand rapidly, the goal of a fully general electronic structure ML model for solids is now within grasp.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.chemmater.1c04252.

Model hyperparameters, predicted DOS curves, error histograms, and extrapolated training size curves (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Victor Fung − *Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States*; orcid.org/0000-0002-3347-6983; Email: fungv@ornl.gov

Authors

P. Ganesh − *Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States*; orcid.org/0000-0002-7170-2902

Bobby G. Sumpter − *Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States*; orcid.org/0000-0001-6341-0355

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.chemmater.1c04252

Author Contributions

V.F. conceived the project and performed the calculations and analysis, and all authors contributed to discussion of the ML-approach, interpretation of the results, and the writing of the manuscript.

Notes

The authors declare no competing financial interest.

The DFT data sets used for training are provided at https://github.com/vxfung/MatDeepLearn_DOS.

The machine learning code used in this work is available at https://github.com/vxfung/MatDeepLearn_DOS.

■ ACKNOWLEDGMENTS

This work (graph neural network prediction (GNN) of DOS for diverse materials) was performed at Oak Ridge National Laboratory's Center for Nanophase Materials Sciences, which is a US Department of Energy Office of Science User Facility. The Center for Understanding and Control of Acid Gas-Induced Evolution of Materials for Energy (UNCAGE-ME), an Energy Frontier Research Center funded by U.S. Department of Energy, Office of Science, Basic Energy Sciences supported extension of GNNs for DOS prediction. V.F. was also supported by a Eugene P. Wigner Fellowship at Oak Ridge National Laboratory. O.R.N.L. is operated by UT-Battelle, LLC., for the U.S. Department of Energy under Contract DEAC05-00OR22725. This research used resources of the National Energy Research Scientific Computing Center, supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

■ REFERENCES

(1) Marzari, N.; Ferretti, A.; Wolverton, C. Electronic-structure methods for materials design. *Nat. Mater.* **2021**, *20*, 736−749.

(2) Curtarolo, S.; Hart, G. L. W.; Nardelli, M. B.; Mingo, N.; Sanvito, S.; Levy, O. The high-throughput highway to computational materials design. *Nat. Mater.* **2013**, *12*, 191−201.

(3) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD). *JOM* **2013**, *65*, 1501−1509.

(4) Choudhary, K.; Garrity, K. F.; Reid, A. C. E.; DeCost, B.; Biacchi, A. J.; Hight Walker, A. R.; Trautt, Z.; Hattrick-Simpers, J.; Kusne, A. G.; Centrone, A.; Davydov, A.; Jiang, J.; Pachter, R.; Cheon, G.; Reed, E.; Agrawal, A.; Qian, X.; Sharma, V.; Zhuang, H.; Kalinin, S. V.; Sumpter, B. G.; Pilania, G.; Acar, P.; Mandal, S.; Haule, K.; Vanderbilt, D.; Rabe, K.; Tavazza, F. The joint automated repository for various integrated simulations (JARVIS) for data-driven materials design. *npj Comput. Mater.* **2020**, *6*, 173.

(5) Curtarolo, S.; Setyawan, W.; Wang, S.; Xue, J.; Yang, K.; Taylor, R. H.; Nelson, L. J.; Hart, G. L. W.; Sanvito, S.; Buongiorno-Nardelli, M.; Mingo, N.; Levy, O. AFLOWLIB. ORG: A distributed materials properties repository from high-throughput ab initio calculations. *Comput. Mater. Sci.* **2012**, *58*, 227−235.

(6) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1*, No. 011002.

(7) Woods-Robinson, R.; Han, Y.; Zhang, H.; Ablekim, T.; Khan, I.; Persson, K. A.; Zakutayev, A. Wide Band Gap Chalcogenide Semiconductors. *Chem. Rev.* **2020**, *120*, 4007−4055.

(8) Nørskov, J. K.; Abild-Pedersen, F.; Studt, F.; Bligaard, T. Density Functional Theory in Surface Chemistry and Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 937−943.

(9) Zhang, L.; Wang, Y.; Lv, J.; Ma, Y. Materials discovery at high pressures. *Nat. Rev. Mater.* **2017**, *2*, 17005.

(10) Park, J. S.; Kim, S.; Xie, Z.; Walsh, A. Point defect engineering in thin-film solar cells. *Nat. Rev. Mater.* **2018**, *3*, 194−210.

(11) Fung, V.; Wu, Z.; Jiang, D.-E. New Bonding Model of Radical Adsorbate on Lattice Oxygen of Perovskites. *J. Phys. Chem. Lett.* **2018**, *9*, 6321−6325.

(12) Singh, A. K.; Montoya, J. H.; Gregoire, J. M.; Persson, K. A. Robust and synthesizable photocatalysts for CO2 reduction: a data-driven materials discovery. *Nat. Commun.* **2019**, *10*, 443.

(13) Yan, Q.; Yu, J.; Suram, S. K.; Zhou, L.; Shinde, A.; Newhouse, P. F.; Chen, W.; Li, G.; Persson, K. A.; Gregoire, J. M.; Neaton, J. B. Solar fuels photoanode materials discovery by integrating high-throughput theory and experiment. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 3040−3043.

(14) Deringer, V. L.; Bernstein, N.; Csányi, G.; Ben Mahmoud, C.; Ceriotti, M.; Wilson, M.; Drabold, D. A.; Elliott, S. R. Origins of structural and electronic transitions in disordered silicon. *Nature* **2021**, *589*, 59−64.

(15) Mannodi-Kanakkithodi, A.; Toriyama, M. Y.; Sen, F. G.; Davis, M. J.; Klie, R. F.; Chan, M. K. Y. Machine-learned impurity level prediction for semiconductors: the example of Cd-based chalcogenides. *npj Comput. Mater.* **2020**, *6*, 39.

(16) Xin, H.; Linic, S. Communications: Exceptions to the d-band model of chemisorption on metal surfaces: The dominant role of repulsion between adsorbate states and metal d-states. *J. Chem. Phys.* **2010**, *132*, 221101.

(17) Andersen, M.; Reuter, K. Adsorption Enthalpies for Catalysis Modeling through Machine-Learned Descriptors. *Acc. Chem. Res.* **2021**, *54*, 2741−2749.

(18) Noh, J.; Back, S.; Kim, J.; Jung, Y. Active learning with non-ab initio input features toward efficient CO2 reduction catalysts. *Chem. Sci.* **2018**, *9*, 5152−5159.

(19) Dickens, C. F.; Montoya, J. H.; Kulkarni, A. R.; Bajdich, M.; Nørskov, J. K. An electronic structure descriptor for oxygen reactivity at metal and metal-oxide surfaces. *Surf. Sci.* **2019**, *681*, 122−129.

(20) Fung, V.; Hu, G.; Sumpter, B. Electronic Band Contraction Induced Low Temperature Methane Activation on Metal Alloys. *J. Mater. Chem. A* **2020**, *8*, 6057−6066.

(21) Wang, S.; Pillai, H. S.; Xin, H. Bayesian learning of chemisorption for bridging the complexity of electronic descriptors. *Nat. Commun.* **2020**, *11*, 6132.

(22) Fung, V.; Hu, G.; Ganesh, P.; Sumpter, B. G. Machine learned features from density of states for accurate adsorption energy prediction. *Nat. Commun.* **2021**, *12*, 88.

(23) Ben Mahmoud, C.; Anelli, A.; Csányi, G.; Ceriotti, M. Learning the electronic density of states in condensed matter. *Phys. Rev. B* **2020**, *102*, No. 235130.

(24) Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **2011**, *134*, No. 074106.

(25) De, S.; Bartók, A. P.; Csányi, G.; Ceriotti, M. Comparing molecules and solids across structural and alchemical space. *Phys. Chem. Chem. Phys.* **2016**, *18*, 13754−13769.

(26) del Rio, B. G.; Kuenneth, C.; Tran, H. D.; Ramprasad, R. An Efficient Deep Learning Scheme To Predict the Electronic Structure of Materials and Molecules: The Example of Graphene-Derived Allotropes. *J. Phys. Chem. A* **2020**, *124*, 9496−9502.

(27) Battaglia, P. W.; Hamrick, J. B.; Bapst, V.; Sanchez-Gonzalez, A.; Zambaldi, V.; Malinowski, M.; Tacchetti, A.; Raposo, D.; Santoro, A.; Faulkner, R. Relational inductive biases, deep learning, and graph networks. 2018, arXiv:1806.01261. arXiv.org e-Print archive, https://arxiv.org/abs/1806.01261 (accessed September 15, 2021).

(28) Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Wang, L.; Li, C.; Sun, M. Graph neural networks: A review of methods and applications. *AI Open* **2020**, *1*, 57−81.

(29) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **2018**, *120*, No. 145301.

(30) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chem. Mater.* **2019**, *31*, 3564−3572.

(31) Fung, V.; Zhang, J.; Juarez, E.; Sumpter, B. G. Benchmarking graph neural networks for materials chemistry. *npj Comput. Mater.* **2021**, *7*, 84.

(32) Back, S.; Yoon, J.; Tian, N.; Zhong, W.; Tran, K.; Ulissi, Z. W. Convolutional Neural Network of Atomic Surface Structures To Predict Binding Energies for High-Throughput Screening of Catalysts. *J. Phys. Chem. Lett.* **2019**, *10*, 4401−4408.

(33) Karniadakis, G. E.; Kevrekidis, I. G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422−440.

(34) Fey, M.; Lenssen, J. E. Fast graph representation learning with PyTorch Geometric. 2019, arXiv:1903.02428. arXiv.org e-Print archive, https://arxiv.org/abs/1903.02428 (accessed September 11, 2021).

(35) Huo, H.; Rupp, M. Unified Representation of Molecules and Crystals for Machine Learning. 2017, arXiv:1704.06439. arXiv.org e-Print archive, https://arxiv.org/abs/1704.06439 (accessed September 10, 2021).

(36) Himanen, L.; Jäger, M. O. J.; Morooka, E. V.; Federici Canova, F.; Ranawat, Y. S.; Gao, D. Z.; Rinke, P.; Foster, A. S. DScribe: Library of descriptors for machine learning in materials science. *Comput. Phys. Commun.* **2020**, *247*, No. 106949.

(37) Mamun, O.; Winther, K. T.; Boes, J. R.; Bligaard, T. A Bayesian framework for adsorption energy prediction on bimetallic alloy catalysts. *npj Comput. Mater.* **2020**, *6*, 177.

(38) Kresse, G.; Furthmuller, 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.

(39) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169.

(40) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(41) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953−17979.

(42) Dudarev, S. L.; Botton, G. A.; Savrasov, S. Y.; Humphreys, C. J.; Sutton, A. P. Electron-Energy-Loss Spectra and the Structural Stability of Nickel Oxide: An LSDA+U Study. *Phys. Rev. B* **1998**, *57*, 1505−1509.

(43) Monkhorst, H. J.; Pack, J. D. Special Points for Brillouin-Zone Integrations. *Phys. Rev. B* **1976**, *13*, 5188−5192.

(44) Jain, A.; Hautier, G.; Moore, C. J.; Ping Ong, S.; Fischer, C. C.; Mueller, T.; Persson, K. A.; Ceder, G. A high-throughput infrastructure for density functional theory calculations. *Comput. Mater. Sci.* **2011**, *50*, 2295−2310.

(45) Mamun, O.; Winther, K. T.; Boes, J. R.; Bligaard, T. High-throughput calculations of catalytic properties of bimetallic alloy surfaces. *Sci. Data* **2019**, *6*, 76.

(46) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine Learning Force Fields. *Chem. Rev.* **2021**, *121*, 10142−10186.

(47) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Quantum chemistry structures and properties of 134 kilo molecules. *Sci. Data* **2014**, *1*, 140022.

(48) Shuaibi, M.; Kolluru, A.; Das, A.; Grover, A.; Sriram, A.; Ulissi, Z.; Zitnick, C. L. Rotation Invariant Graph Neural Networks using Spin Convolutions. 2021, arXiv:2106.09575. arXiv.org e-Print archive, https://arxiv.org/abs/2106.09575 (accessed September 20, 2021).

(49) Godwin, J.; Schaarschmidt, M.; Gaunt, A.; Sanchez-Gonzalez, A.; Rubanova, Y.; Veličković, P.; Kirkpatrick, J.; Battaglia, P. Simple GNN Regularisation for 3D Molecular Property Prediction & Beyond. 2021, arXiv:2106.07971v2. arXiv.org e-Print archive, https://arxiv.org/abs/2106.07971 (accessed March 20, 2022).

(50) Choudhary, K.; DeCost, B. Atomistic Line Graph Neural Network for improved materials property predictions. *npj Comput. Mater.* **2021**, *7*, 185.

(51) Bang, K.; Yeo, B. C.; Kim, D.; Han, S. S.; Lee, H. M. Accelerated mapping of electronic density of states patterns of metallic nanoparticles via machine-learning. *Sci. Rep.* **2021**, *11*, 11604.

(52) Yeo, B. C.; Kim, D.; Kim, C.; Han, S. S. Pattern Learning Electronic Density of States. *Sci. Rep.* **2019**, *9*, 5879.

(53) Kong, S.; Ricci, F.; Guevarra, D.; Neaton, J. B.; Gomes, C. P.; Gregoire, J. M. Density of states prediction for materials discovery via contrastive learning from probabilistic embeddings. *Nat. Commun.* **2022**, *13*, 949.

(54) Kaundinya, P. R.; Choudhary, K.; Kalidindi, S. R. Prediction of the Electron Density of States for Crystalline Compounds with Atomistic Line Graph Neural Networks (ALIGNN). *JOM* **2022**, *74*, 1395−1405.

(55) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Big Data Meets Quantum Chemistry Approximations: The Δ-Machine Learning Approach. *J. Chem. Theory Comput.* **2015**, *11*, 2087−2096.

(56) Chen, Z.; Andrejevic, N.; Smidt, T.; Ding, Z.; Xu, Q.; Chi, Y.-T.; Nguyen, Q. T.; Alatas, A.; Kong, J.; Li, M. Direct Prediction of Phonon Density of States With Euclidean Neural Networks. *Adv. Sci.* **2021**, *8*, No. 2004214.