

Predicting HSE band gaps from PBE charge densities via neural network functionals

Levi C Lentz and Alexie M Kolpak[✉]

Massachusetts Institute of Technology, Mechanical Engineering, Cambridge, MA 02139,
United States of America

E-mail: llentz@mit.edu and kolpak@mit.edu

Received 20 August 2019, revised 27 November 2019

Accepted for publication 5 December 2019

Published 9 January 2020



Abstract

Density functional theory (DFT) has become a standard method for *ab initio* calculations of material properties. However, it has a number of shortcomings, particularly in predicting key properties, such as band gap and optical spectra, which are dependent on excited states. To treat such properties, more accurate approaches such as GW or DFT with hybrid functionals (including HSE, PBE0, and B3LYP, to name a few) can be employed; however, these approaches are unfeasible for many large and/or complex systems due to their high computational cost and large memory requirements. In this work, we investigate the ability to train neural networks of the traditional DFT charge density computed with a standard PBE functional to accurately predict HSE band gaps. We show that a single network PBE charge density functional can predict the HSE band gap of seven different materials—silicon, gallium arsenide, molybdenum disulfide, germanium, tin phosphate, titanium phosphate, and zirconium phosphate—under a wide variety of conditions with an RSME of 172.6 meV, which is 34% better accuracy than standard regression between the PBE and HSE band gaps. This approach, which, in principle, can be used to map PBE charge densities to band gaps or other properties computed with any higher accuracy method, has the potential to decrease computational costs, increase prediction accuracy, and enable accurate high-throughput screening for a wide variety of complex materials systems.

Keywords: density functional theory, band gap, machine learning

(Some figures may appear in colour only in the online journal)

Density functional theory (DFT) has become the standard for studying periodic systems for a wide range of materials properties. Originally proposed in the 1960s [1, 2], this formalism states that material properties can be described as a pure functional of the charge density, eliminating the need to directly treat all the electrons in the system. This finding drastically reduced the computational costs associated with simulating realistic systems to study their unique properties.

While DFT is promising for a variety of reasons, there are several shortcomings of the theory. One of the most significant shortcomings of DFT is that large systems are difficult to study. This is because some of the terms within DFT formalism are not known as a pure functional of the charge density. One example of this is the kinetic energy term, requiring explicit treatment of the Kohn-Sham orbitals, drastically increasing the computational cost of DFT methods. Work in

orbital-free DFT [3] has been proposed to decrease this reliance and is a promising way to achieve simulations of tens of thousands of atoms within DFT.

Arising from the lack of an exact functional, a second area of concern is in predicting certain materials properties accurately. While DFT is useful for predicting formation energies, potential energy surfaces, total energies, and ground state atomic structures, it performs poorly for predicting properties that are inherently dependent on excited states. Commonly used DFT exchange-correlation approximations, such as the local density approximation (LDA) [2] and generalized gradient approximation (GGA) [4], are notoriously inaccurate at predicting key properties such as the band gap. This is generally thought to be caused by the inaccurate treatment of the self-interaction term of the electrons [5]. Higher-order methods such as the GW approximation (GWA) [6] are capable

of calculating the band gap with much higher accuracy, but at high computational cost and only for small systems.

One alternative is to employ hybrid exchange such as Heyd–Scuseria–Ernzerhof (HSE) [7] or Becke, three-parameter, Lee–Yang–Parr (B3LYP) [8, 9]. These methods modify the Perdew–Burke–Ernzerhof (PBE) exchange energy with portions of the exact exchange, parameterized to ensure agreement with experiment for specific properties over a specific data set of materials. While these methods obtain better results, they are also too computationally expensive to employ for high-throughput screening or on very large systems. Decreasing the computational cost of these methods would enable researchers to apply these to a wider variety of systems not currently possible with standard techniques.

In this work, we explore using machine learning methods, specifically neural networks (NNs), to decrease the computational cost of higher accuracy calculations, using HSE calculations as an example target higher-order method. We exploit this machinery to map PBE charge densities to the HSE band gap by effectively learning the functional of the PBE charge density required to predict the HSE band gap. While this method does not directly correct the PBE charge density to account for the self-interaction term, it is a powerful way to map computationally cheap methods such as PBE to relatively expensive methods such as HSE through training on small datasets. In this work we demonstrate the power of NNs to learn this functional form over a wide range of systems and species; expansion of this work to other material properties could enable NNs to augment traditional DFT methods to both lower computational costs and increase accuracy. Uniquely, this method does not rely on any transformation of the input dataset; the network functional that is learned exists within the charge-density space, potentially enabling analysis of the NN functional to better understand how higher-order methods correct inaccuracies in methods such as PBE.

1. Approach

In recent years, there has been an explosion of machine learning methods into the materials science community. Examples include Gaussian approximation potentials [10], Gaussian process regression [11], ordinary least squares with feature engineering [12], smooth overlap of atomistic potentials (SOAP) [13], crystal graph convolutional networks [14], NN machine learning methods [15–17], and atomistic fingerprinting [18–23]. These methods seek to map atomic positions (both molecules and periodic systems) to a DFT-property of interest. While these methods are powerful for their purpose, they do not directly use charge-densities in their learning, preventing the study of charge-density relationships. In our previous work [24], we proposed a method of directly training on the charge density, the fundamental feature of DFT.

Several other methods to learn information from the charge density, including predicting densities from Hodenberg–Kohn Fourier components [25] and using convolution networks to learn the structure in charge densities [26], have been proposed. Unlike our approach, however, these methods rely

heavily on transforming and augmenting the input charge-density space, making direct comparison of the charge density to the machine-learned structure difficult. While these methods show great promise, our previously proposed method [24] has an advantage in that it specifically trains to the functional of the charge density, implicitly preserving all rotational and translational symmetries of the system, thereby dramatically decreasing the amount of input data required. Further, training directly to the kernel of the charge density removes the need to augment the charge-density input space, enabling researchers to directly compare the learned NN charge-density functionals to existing functionals. In brief, this network learns the following relationship

$$\Theta[\rho(\vec{r})] = \int f(\rho(\vec{x}_1), \rho(\vec{x}_2), \dots, \rho(\vec{x}_n)) d^n V, \quad (1)$$

where Θ is the functional, ρ is the charge-density at a point \vec{x} , and f is the kernel of the functional. Training is performed in integration-space, implying that NN sits at the integrand of the functional, effectively learning f . A more complete breakdown of training can be found in [24].

In this article we build on our previous work on molecular systems by demonstrating the power of NN functionals to learn a functional for a wide variety of complex solid state systems. We examine several low, medium, and large band gap semiconductors of interest to optoelectronic applications, including traditional semiconductors such as silicon (Si), germanium (Ge), gallium arsenide (GaAs), and molybdenum disulfide (MoS₂). Additionally, we consider three transition metal phosphates, tin phosphate (SnPhos [27]), zirconium phosphate (ZrPhos [28]), and titanium phosphate (TiPhos [29]), chosen for their large band gaps, significantly different crystal structures relative to the other materials in the training set, and because we have previously worked extensively with these materials due to their promising properties for forming heterostructures and band gap engineering [30, 31].

Initial bulk structures were first fully relaxed using DFT with the Perdew–Burke–Ernzerhof (PBE) exchange correlation functional. Relaxations were performed with high energy cutoff and strict convergence criteria, ensuring we find the complete ground state structures. Further details of the calculations are described in the methods section. Relaxed structures were used as seed structures for *ab initio* molecular dynamics (AIMD) calculations to generate a variety of structures for the training, testing, and validation sets. To obtain a wide range of structures, AIMD calculations were performed with significantly lower convergence criteria, generating both realistic and unrealistic structures with a significant range of band gap, as represented in figure 1. The results of the AIMD simulation are shown in figure 1. In the AIMD simulations, we used a time step of 1 femtosecond through all simulations with velocity rescaling performed every 40–100 time steps. The simulated temperatures ranged from 300–600 K.

After each AIMD simulation, the structure, energy, and band gap (from the AIMD calculation) for each material at each time step were collected. From the resulting structures, we selected 50 structures uniformly along the band gap dimension for each material studied. As unrealistic structures were

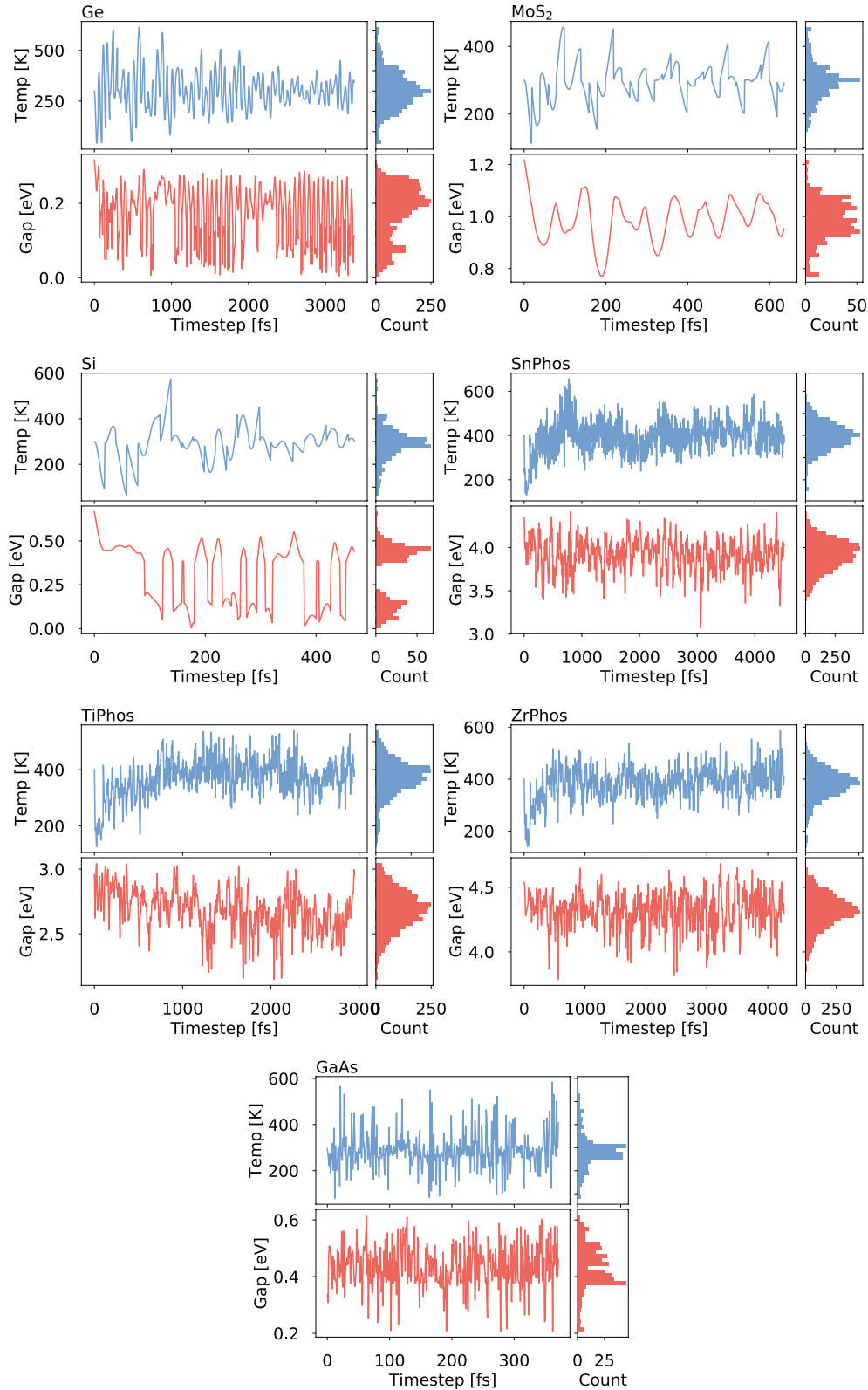


Figure 1. Representative AIMD data used to generate the structures in this work. The AIMD calculations were performed purely to generate a wide range of structures and corresponding band gaps.

present within this dataset, structures outside of three standard deviations of the band gap were not considered. This resulted in 350 total bulk structures. While we used uniform sampling to sample over the range of AIMD structures, we note that other

sampling methods may better approximate the structure-band gap relationship, potentially reducing the number of samples needed for the NN fit. We then re-ran these extracted structures with high-convergent SCF PBE calculations to generate

Table 1. Tabulated information on materials, number of atoms, dopant species, MSD for both the sample data and the AIMD data, PBE band gap range, and HSE band gap range. Sampling was performed based on the band gap; however, the resulting MSD for the sample set is in good agreement with the entire AIMD set.

Phase	Atoms	Dopant (Site)	Sample MSD (\AA^2)	AIMD MSD (\AA^2)	PBE Gap (eV)	HSE Gap (eV)
MoS ₂	24	Cr (Mo), V (Mo), Nb (Mo), N (S)	0.047	0.051	0.31–0.92	0.16–1.4
Si	8	Fe	0.019	0.018	0–0.56	0.64–1.17
GaAs	8	—	0.139	0.137	0.21–0.62	0.73–1.54
Ge	8	—	0.013	0.012	0.01–0.23	0.08–0.87
SnPhos	26	—	0.031	0.032	3.94–4.34	5.43–6.30
TiPhos	26	—	0.248	0.246	2.34–2.98	3.87–4.57
ZrPhos	26	—	0.09	0.088	4.01–4.58	5.66–6.28

PBE charge densities. We doped MoS₂ and Si structures with several dopants to generate a range of *p*- and *n*-type character [32]; dopant atoms were placed at random atomic positions within the respective 50-sample sets. The final dataset was then run with the Heyd–Scuseria–Ernzerhof (HSE) exchange correlation to generate a more accurate band gap. We note that the same procedure could be used with even higher accuracy methods such as GW.

The final dataset is tabulated in table 1, which shows the bulk materials, dopants, and mean squared displacement (MSD) for each material, as well as the PBE and HSE band gap ranges. Even though sampling was performed by only considering the PBE bandgap of the AIMD run, the sample MSD is similar to the AIMD MSD for all phases, indicating that the sampled set is indicative of the AIMD simulation populations. During training, only non-metallic systems were considered, limiting the overall number of systems available to train. In total, 538 PBE charge density–HSE band gap pairs were used for training.

Finally, the real-space PBE charge densities were the input to a neural network (NN) with the target of the network being the HSE band gap. Training of the NN functional was performed using the materials property prophet (PROPhet), as described in [24].

2. Results and discussion

The majority of this work centers around the results and implications of the NN training. However, it is worth examining if a powerful method such as a NN can outperform simpler methods. As we have a wide range of HSE calculation examples, it is informative to compare all NN results to a linear regression model. A simple linear regression has the form $E_{\text{gap}}^{\text{HSE}} = m \cdot E_{\text{gap}}^{\text{PBE}} + b$, where the *m* and *b* hyperparameters are found by using standard linear regression techniques. While other, more complex, machine learning methods such as support vector machines could be deployed on this dataset, these would require feature engineering and selection, reducing the generalizability of the learned model. As it is commonly assumed that the PBE exchange correlation systematically under-predicts the band gap of a system, being able to outperform simple linear regression would be indicative that NN functionals would be able to succeed in complex, and

non-linear, examples. Further, both regression and NN potentials do not transform the input space, allowing for direct comparison.

To compare these methods directly we employ two common scores. As with other machine learning methods, we measure our results against the root-mean-square deviation, $\text{RMSE} = (n^{-1} \sum_i (x_{i,t} - x_{i,p})^2)^{1/2}$, where *i* ranges over the number of samples *n*, $x_{i,t}$ is the target for *i*, and $x_{i,p}$ is the prediction for *i*. During training, the gradient is calculated from the RMSE. For comparisons to other works, we also compare our results via the mean absolute error, $\text{MAE} = n^{-1} \sum_i |x_{i,t} - x_{i,p}|$, where all variables on the right hand side have the same nomenclature as the RMSE.

The dataset along with linear regression results are shown graphically in figure 2. We explored using higher-order fits for this dataset, however these did not yield more accurate regression fits, likely because of the strong linearity present in this dataset. We only consider non-metallic systems as here we are concerned with accurately predicting the band gap of insulating materials. In addition, it is unintuitive, from a machine learning standpoint, to represent both metallic and insulating structures in a single functional. The overall RMSE and MAE for a single linear regression are 360.3 meV and 274.7 meV, respectively.

Training was performed in a typical 80%–10%–10% split into training, validation, and test datasets, respectively. This means that for each dopant within a given phase of material, 80% of the data was selected, randomly, for training. The remaining 20% of the data was randomly split into validation and test datasets. During training, we calculate the RMSE for the validation dataset at each epoch; the functional that minimized the error on the validation data was the functional used on the test dataset. This is required to guarantee the generality of the functional, as training and validation data are not independent of the network.

A common problem of NNs are their sensitivity to their initial starting weights. This problem is further exasperated on datasets such as this where there is a high dimensionality to the input space and low number of training samples. To help address this, we employ a committee of NNs [33, 34] that are averaged to a final output, preventing an anomalous single fit from characterizing this method’s performance. In this regard, we train five models, each with randomized starting weights,

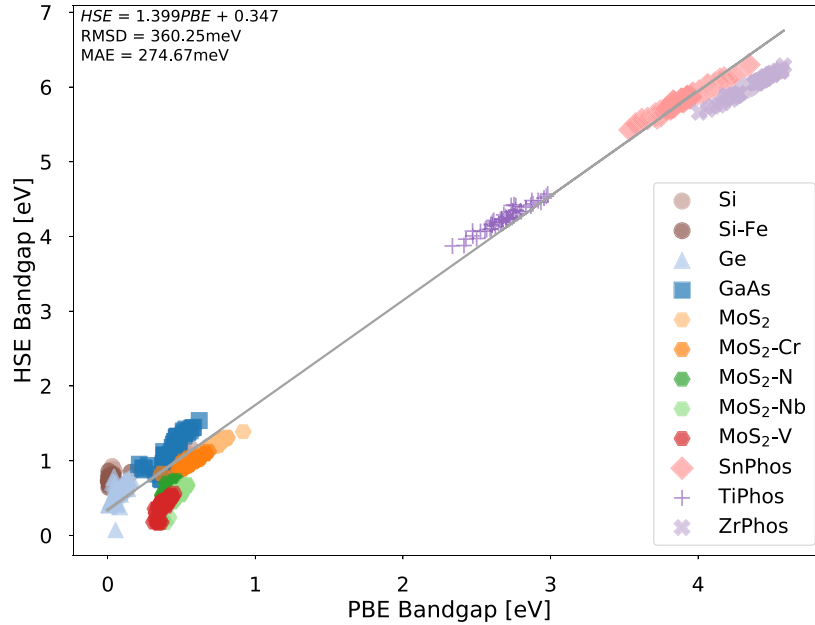


Figure 2. HSE band gap versus PBE band gap. From the sampled systems, we placed dopants at random atomic sites in the simulation cell. The RMSE for this regression is in the upper left of the subplot. The range of systems make it difficult for linear regression to learn the relationship between the PBE band gap and the HSE band gap.

randomized train, test, and validate sets, but with identical network topologies. A proper committee would require a range of topologies, however the computational complexity of the data generation and network training precludes studying deep networks within this work. For comparison to the linear regression methods, the same training methodology is used: randomized training data as well as random starting hyperparameters with the output of these models averaged over all five regression fits.

Charge density sizes ranged from a minimum of approximately 300000 fast Fourier transform (FFT) grid points (for Si) to a maximum of 1.38×10^6 FFT grid points (for ZrPhos). While the network can interpret the size differences in the charge densities, the computational cost of training is large and grows with the number of data samples; it is possible to run into memory issues associated with training on a large number of large charge densities. In this work, we employ downsampling as a way to decrease the complexity of the charge density. This method, standardized in our previous work [24], decreases the charge density by spatially averaging the density on the FFT grid. For example, a downsample of 1, or parity, is effectively directly to the charge density. A downsample of 2 decreases the charge density size by a factor of 8 (or by halving the charge density in each of the principle directions), by averaging each point in the initial charge density over the nearest neighbor of the FFT grid. FFT points are not double counted in this method: if a downsample of 2 is employed a stride along the FFT grid of 2 is also used, implying each FFT grid point is spatially averaged only once. By extension, a downsample of 3 would decrease the charge density by a factor of 27, and so forth.

For systems with low variability in the charge density relative to the property of interest, we can employ a large

downsample to decrease the computational cost of the training. Likewise, for systems with high variability in their charge density, a lower, or no downsample, is desired. In single-phase systems with no dopants, a high downsample can be used and achieve good results. In this work, we converged the downsample required to achieve optimal results: a downsample of 2 is used for all NN functionals.

All networks studied in this work have an input layer (with a single input neuron), a hidden layer, followed by an integration layer as the output (to ensure the network is the kernel of the charge density). All hidden neurons have hyperbolic tangent (tanh) activation functions while input and output nodes have linear activation functions. We feed the charge density into the network sequentially, looping over the FFT grid in the same order for all data samples. The output is stored for each input and finally integrated. Minimization of the network is performed on the error associated with integration. This process enables the network to handle charge densities of varying sizes as the single input neuron and integration layer removes any size dependence imposed on the network by the charge density.

Several different network topologies were considered, including a single hidden layer, two hidden layers, and three hidden layers. Given the size of the dataset, we find that a single hidden layer is able to outperform more complex networks. A likely reason for this is due to the kernel nature of the network. This formalism means that training happens in integration space, enabling the network to implicitly account for the spatial dependence in the charge density. A more complex network would be required if training relied on understanding the overall structure of the charge density.

All the phases represented in figure 2 were used to fit five models with 150 neurons in the single hidden layer. The larger

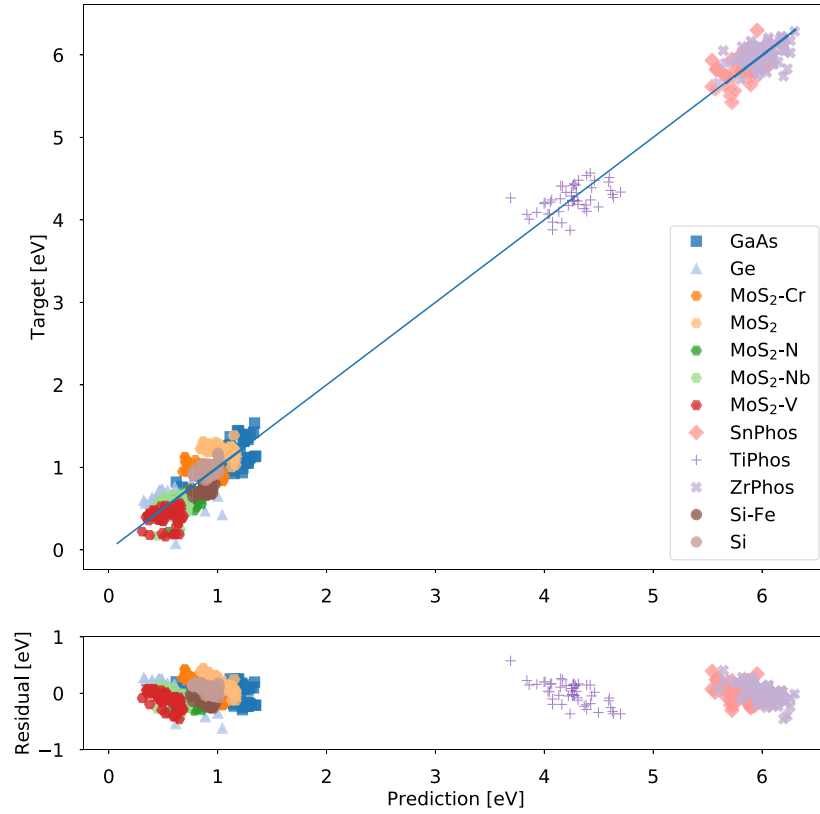


Figure 3. Average output of five 150-neuron single-layer network functionals. Average RMS error of 172.6 meV compared to the linear regression result of 263.1 meV. This network is able to learn the functional that maps PBE charge density to HSE band gap with high accuracy over a range of systems and a large range of band gaps.

layer was possible because of the large amount of data available to the network; the large single layer effectively behaved as a classification layer, learning the bandgap relationship for each phase (see supplementary information, available online at stacks.iop.org/JPhysCM/32/155901/mmedia), allowing accurate prediction for a wide range of systems and dopants. Qualitative fit information is shown in figure 3, showing good agreement over all system classes. The tabulated results are shown in table 2, where the network outperforms linear regression for both RMSE and MAE over all systems and material classes. The distribution of errors are shown in figure 4 indicating the NN is more precise and accurate than the regression fit across all phases. Errors are generally centered around zero, indicating that the network is not sacrificing accuracy in certain phases to achieve a lower error, as the regression fit does.

As a final test of the ability of the model to generalize, several phases were run again within AIMD at 600 K. Specifically, Si, GaAs, and MoS₂ were examined with structures selected from the AIMD trajectory in the same fashion as before, however approximately 10 systems/phase were examined. Similar dopants were placed in respective phases. Shown in table 3, the NN functional outperforms the regression fit on 600 K trajectories. However, we see a slightly higher error than before.

Table 2. Material-specific statistics on the performance of the NN and the regression analysis. The NN is able to outperform the regression model for all phases, and generally by a significant amount. All units are meV.

Phase	RMSE		MAE	
	Regression	NN	Regression	NN
GaAs	303.68	145.89	279.46	122.97
Ge	264.25	206.78	253.40	164.75
MoS ₂	257.10	182.41	206.88	142.56
Si	289.22	120.45	231.14	98.94
SnPhos	181.10	151.01	167.70	119.78
TiPhos	280.98	196.38	274.48	160.07
ZrPhos	258.43	169.33	254.87	132.87

Using principal component analysis to visually analyze a low-dimensional representation of the charge densities, we find that these 600 K systems are outside the bounds of the training set (see supplementary information). While the network is still outperforms the regression, it is likely that multiple MD trajectories would have led to a more robust training set, allowing the network to learn a more general representation of the functional that transforms the PBE charge density to the HSE band gap.

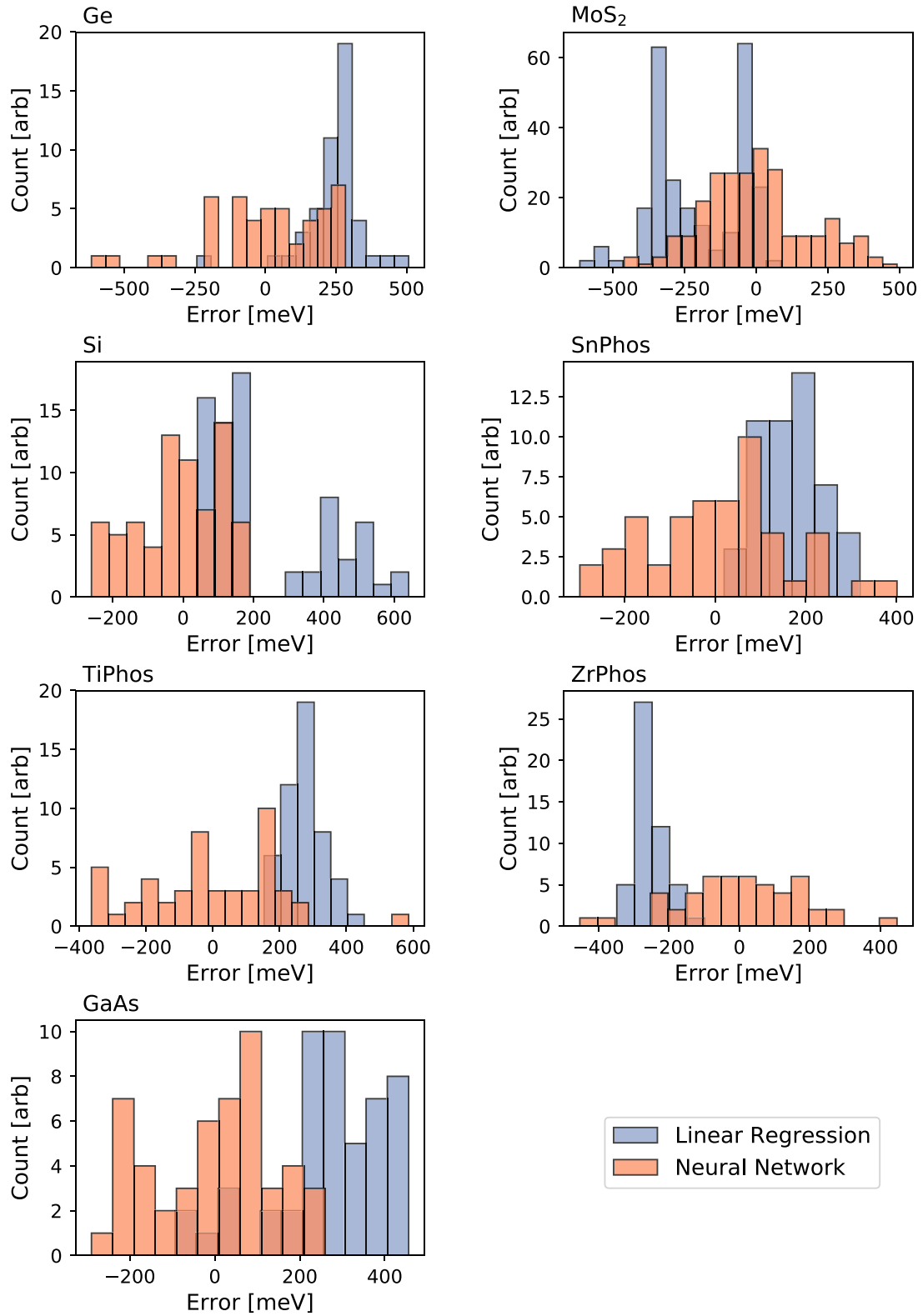


Figure 4. Errors associated with NN fit and linear regression fit. While there are outliers for both fits, generally the NN fit is centered around a zero-mean while the regression fit's mean is displaced from zero. This indicates that the NN fit is generalizing to all the phases, whereas the regression fit generally has a systemic error that varies per phase.

Table 3. Validation of the network and regression models on data selected from 600 K trajectories. Here we see that while overall errors are higher, NN functionals are able to significantly outperform regression analysis. All units in meV.

Phase	RMSE		MAE	
	Regression	NN	Regression	NN
GaAs	371.58	236.79	360.39	179.34
MoS ₂	287.21	244.99	213.11	208.63
Si	311.92	137.94	250.17	107.87

3. Conclusion

In this work, we have demonstrated a potential application for NN functionals: to address the shortcomings of PBE functionals in predicting the accurate band gap of materials. We have demonstrated the ability to use PBE charge densities as input data to a NN to learn the relationship between the PBE charge density and the HSE band gap. We have shown that NNs can outperform simple statistical methods such as linear regression. Outperforming linear regression here is an indication that NN functionals could drastically increase the accuracy of modeling non-linear relationships. This method could aide in learning relationships where a functional relationship is completely unknown. This could, potentially, be a boon to the computational community as this would significantly decrease the computational cost of estimating properties dependent on exact DFT methods.

4. Methods

In this work, all calculations were performed with Quantum Espresso [35]. Norm conserving pseudopotentials were selected from the Pseudo Dojo repository [36, 37] to guarantee consistency across the material species types. Data generated for training were generated by performing *ab initio* molecular dynamics over a temperature range of 300–600 K as described in the text. During data generation, timesteps were 1 fs with temperature rescaling every 40–100 timesteps; *ab initio* MD was performed only to generate reasonable structures rather than to study the strict temperature dependent nature of the band gap. We ran AIMD calculations with relaxed cutoff criteria, specifically with *k*-point meshes coarser than the actual SCF calculations. During AIMD, we employed charge density interpolation between timesteps to increase the speed of the SCF calculation within the AIMD run.

During AIMD calculations the Perdew–Burke–Ernzerhof [4] (PBE) exchange correlation was used. After we extracted structures from the AIMD run, we generated fully converged PBE charge densities for these selected structures. The cutoff and *k*-point mesh were chosen to converge forces and total energies within 10 meV Å⁻¹ and 20 meV, respectively. This represented planewave cutoffs of 110 Ry, 90 Ry, 120 Ry, 120 Ry, 140 Ry, 140 Ry, and 140 Ry for MoS₂, Si, GaAs, Ge, SnPhos, TiPhos, and ZrPhos respectively. Under these same criteria, the *k*-mesh was 3 × 3 × 3, 5 × 5 × 5, 5 × 5 × 5, 6 × 6 × 6, 2 × 3 × 2, 2 × 3 × 2, and 2 × 3 × 2, respectively. To

accurately calculate the band gap for these structures, these were then run within the Heyd–Scuseria–Ernzerhof [7] (HSE) formalism. We used the same energy cutoff with the *k*-mesh chosen to converge the band gap. This led to a uniform *k*-mesh for the HSE calculations of 3 × 3 × 1, 4 × 4 × 4, 2 × 2 × 2, 2 × 2 × 2, 1 × 1 × 1, 1 × 1 × 1, and 1 × 1 × 1 for MoS₂, Si, GaAs, Ge, SnPhos, and ZrPhos respectively.

Acknowledgments

Research supported as part of S3TEC, an Energy Frontier Research Center (EFRC) funded by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), under Award DE-SC0001299/DE-FG02-09ER46577.

Author contributions

LCL generated all data and performed NN modeling. LCL and AMK performed the analysis and manuscript preparation.

Competing interests

The authors declare no competing interests.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID iDs

Alexie M Kolpak  <https://orcid.org/0000-0002-4347-0139>

References

- [1] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
- [2] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
- [3] Yip S 2005 *Handbook of Materials Modeling* (Berlin: Springer) pp 137–48
- [4] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
- [5] Perdew J P 1985 *Int. J. Quantum Chem.* **28** 497
- [6] Aryasetiawan F and Gunnarsson O 1998 *Rep. Prog. Phys.* **61** 237
- [7] Heyd J, Scuseria G E and Ernzerhof M 2003 *Chem. Phys.* **118** 8207
- [8] Stephens P, Devlin F, Chabalowski C and Frisch M J 1994 *J. Phys. Chem.* **98** 11623
- [9] Kim K and Jordan K 1994 *J. Phys. Chem.* **98** 10089
- [10] Bartók A P, Payne M C, Kondor R and Csányi G 2010 *Phys. Rev. Lett.* **104** 136403
- [11] Piliñia G, Gubernatis J E and Lookman T 2017 *Comput. Mater. Sci.* **129** 156
- [12] Lee J, Seko A, Shitara K, Nakayama K and Tanaka I 2016 *Phys. Rev. B* **93** 115104
- [13] Bartók A P, Kondor R and Csányi G 2013 *Phys. Rev. B* **87** 184115
- [14] Xie T and Grossman J C 2018 *Phys. Rev. Lett.* **120** 145301

- [15] Schleder G R, Padilha A C, Acosta C M, Costa M and Fazzio A 2019 *J. Phys.: Mater.* **2** 032001
- [16] Nemnes G, Mitran T and Manolescu A 2019 *J. Nanomaterials* **2019** 6960787
- [17] Zhuo Y, Mansouri Tehrani A and Brgoch J 2018 *J. Phys. Chem. Lett.* **9** 1668
- [18] Tang Y H, Zhang D and Karniadakis G E 2018 *J. Chem. Phys.* **148** 034101
- [19] Burden F R, Polley M J and Winkler D A 2009 *J. Chem. Inf. Model.* **49** 710
- [20] Huan T D, Mannodi-Kanakkithodi A and Ramprasad R 2015 *Phys. Rev. B* **92** 014106
- [21] Behler J and Parrinello M 2007 *Phys. Rev. Lett.* **98** 146401
- [22] Behler J 2011 *Phys. Chem. Chem. Phys.* **13** 17930
- [23] Behler J 2014 *J. Phys.: Condens. Matter* **26** 183001
- [24] Kolb B, Lentz L C and Kolpak A M 2017 *Sci. Rep.* **7** 1192
- [25] Brockherde F, Vogt L, Li L, Tuckerman M E, Burke K and Müller K R 2017 *Nat. Commun.* **8** 872
- [26] Kajita S, Ohba N, Jinnouchi R and Asahi R 2017 *Sci. Rep.* **7** 16991
- [27] Bruque S, Aranda M A, Losilla E R, Olivera-Pastor P and Maireles-Torres P 1995 *Inorg. Chem.* **34** 893
- [28] Millini R, Perego G, Costantino U and Marmottini F 1993 *Microporous Mater.* **2** 41
- [29] Salvadó M, Pertierra P, García-Granda S, García J, Rodríguez J and Fernández-Díaz M 1996 *Acta Crystallogr. B* **52** 896
- [30] Lentz L C, Kolb B and Kolpak A M 2016 *Phys. Chem. Chem. Phys.* **18** 14122
- [31] Lentz L C and Kolpak A M 2017 *Sci. Rep.* **7** 1248
- [32] Dolui K, Rungger I, Das Pemmaraju C and Sanvito S 2013 *Phys. Rev. B* **88** 075420
- [33] Guo J J and Luh P B 2004 *IEEE Trans. Power Syst.* **19** 1867
- [34] Fernandes A M, Utkin A B, Lavrov A V and Vilar R M 2004 *Pattern Recognit.* **37** 2039
- [35] Giannozzi P *et al* 2009 *J. Phys.: Condens. Matter* **21** 395502
- [36] Lejaeghere K *et al* 2016 *Science* **351** aad3000
- [37] Hamann D R 2013 *Phys. Rev. B* **88** 085117