Multi-fidelity Graph Networks for Machine Learning the Experimental Properties of Ordered and Disordered Materials

Chi Chen, Yunxing Zuo, Weike Ye, Xiangguo Li, and Shyue Ping Ong*

Department of NanoEngineering, University of California San Diego, CA, USA

E-mail: ongsp@eng.ucsd.edu

Abstract

Predicting the properties of a material from the arrangement of its atoms is a fundamental goal in materials science. In recent years, machine learning (ML) on ab initio calculations has emerged as a new paradigm to provide rapid predictions of materials properties across vast chemical spaces. However, the performance of ML models are determined by the quantity and quality of data, which tend to be inversely correlated with each other. Here, we develop multi-fidelity materials graph networks (MFGNet) to transcend this trade-off to achieve accurate predictions of the experimental band gaps of ordered and disordered materials to within 0.3-0.5 eV. We show that the inclusion of low-fidelity Perdew-Burke-Ernzerhof band gaps significantly enhances the resolution of latent structural features in materials graph representations, leading to 22-45% decrease in the mean absolute errors of high-fidelity computed and experimental band gap predictions with an order of magnitude smaller data sizes. Further, MFGNet
models can be readily extended to predict the band gaps of disordered crystals to excellent agreement with experiments, addressing a major gap in the computational prediction of materials properties.

In silico predictions of the properties of materials can most reliably be carried out using ab initio calculations. However, their high computational expense and poor scalability have limited their application to mostly to materials containing < 1000 atoms without site disorder. Further, a rule of thumb is that the more accurate the ab initio method, the higher the computational expense and the poorer the scalability. It is therefore no surprise that supervised machine learning (ML) of ab initio calculations has garnered substantial interest as a means to develop efficient surrogate models for materials property predictions. State-of-the-art ML models encode structural information (e.g., as graphs or local environmental features) in addition to composition information, allowing them to distinguish between polymorphs that may have vastly different properties.

Most frustratingly, while building ML models from high-accuracy calculations or experiments would yield the greatest value, obtaining sufficient data to reliably train such models is also the most challenging. For example, the number of PBE calculations in large, public databases such as the Materials Project and Open Quantum Materials Database (OQMD) is on the order of $10^6$, while the number of Heyd-Scuseria-Ernzerhof (HSE) calculations is at least two orders of magnitude fewer. Similarly, while B3LYP/PBE calculations are available for millions of molecules, “gold standard” CCSD(T) calculations are only available for perhaps thousands of small molecules. Even fewer are the number of high-quality experimental data points.

A potential approach to address this challenge is through multi-fidelity models, i.e., models that combine low-fidelity data with high-fidelity data. In the handful of previous works utilizing this approach in ML of materials properties, all are two-fidelity models utilizing a co-kridding approach, which assumes an approximately linear relationship between targets of different fidelity. The training of co-kridding models scales as $O(N^3)$ (where $N$ is the num-
ber of data points), which becomes prohibitively expensive when \( N \) exceeds 10,000. Further, these efforts have been limited to specific properties of single structure prototypes.

Here, we develop multi-fidelity materials graph networks (MFGNet) as a generalized framework for materials property prediction across computational methodologies and experiments, as shown in Fig. 1. Graph networks are a general, composable deep learning framework that supports both relational reasoning and combinatorial generalization.\(^{3,4}\) Previously, the current authors have shown that materials graph network models can significantly outperform prior ML models in predicting the properties of both molecules and crystals.\(^{4}\)

The starting point is a natural graph representation of a material, where the atoms are the nodes and the bond between them are the edges. In this work, the input atomic attribute is simply the atomic number of the element passed to a trainable embedding matrix. The bond attribute is the Gaussian-expanded distance. The state attribute vector provides a portal for structural-independent features to be incorporated into the model. Here, the fidelity level (e.g., computational methods or experimental) is encoded as an integer and passed to a trainable embedding matrix to form the input state attributes. An MFGNet model is built by concatenating a series of graph convolutional layers that sequentially exchanges information between atoms, bonds and the state vector. In the final step, the latent features in the output graph is read out and passed into a neural network to arrive at a property prediction. Further details are available in the Methods section.

We have selected the prediction of the band gap (\( E_g \)) of crystals as the target problem because of its great importance in a broad range of technological applications, including photovoltaics, solar water splitting, etc., as well as the availability of data of multiple fidelities. The low fidelity (low-fi) dataset comprise 52,348 PBE band gaps from the Materials Project.\(^{13}\) The high fidelity (high-fi) computed datasets comprise 2,290 Gritsenko-Leeuwen-Lenthe-Baerends with solid correction (GLLB-SC),\(^{20–22}\) 472 strongly constrained and appropriately normed (SCAN)\(^{23,24}\) and 6,030 Heyd-Scuseria-Ernzerhof (HSE)\(^{15,25}\) band gaps. Experimental band gaps for 2,703 ordered crystals and 278 disordered crystals\(^{26}\) are consid-
Figure 1: **Multi-fidelity graph networks.** a, Representation of a material in a graph network model, with atoms as the nodes, bonds as the edges coupled with a structure-independent global state. The atomic number of the element is the input atomic feature and is embedded to a length-16 feature vector. The Gaussian expanded distance is used as the bond feature vector. In MFGNet, the fidelity of each data (computational method or experiment) is encoded as an integer and embedded into a length-16 global state feature vector. b, A MFGNet model is constructed by stacking graph convolution layers. In each graph convolution layer, sequential updates of atomic, bond and state features are performed using information from connected neighbors in the graph. The output graph in the last layer is then readout and processed in a neural network to arrive at the final prediction.
ered a separate high fidelity dataset. The computationally least expensive PBE functional is well-known to systematically underestimate the band gap, and the high-fi functionals correct this to varying degrees. The data within each fidelity was randomly split into 80% training, 10% validation and 10% test data. Model training was carried out using the training and validation sets for a maximum of 1,500 epochs, and final model performances were evaluated on the test set. This process was repeated five times to obtain an averaged mean absolute error (MAE) with error bars.

MFGNet models for the band gaps of ordered crystals were first developed for each fidelity in isolation, i.e., single-fidelity or 1-fi models. The MAE of the 1-fi MFGNet models (Fig. 2a) is related to the data size as well as the mean absolute deviation (MAD, see Table S1) within each dataset. The PBE dataset is the largest with a small MAD, and the 1-fi PBE models have the smallest average MAE of 0.27 eV. The average MAEs for the computed 1-fi models increases in the order of $\text{PBE} < \text{HSE} < \text{GLLB-SC} < \text{SCAN}$, in inverse order to the dataset size. The lower MAE of the experimental 1-fi model compared to the HSE 1-fi model despite having a smaller data set size may be attributed to the relatively large fraction of metals (with zero band gap) in that dataset, which leads to smaller MAD.

Multi-fidelity models utilizing the large PBE dataset with data from other fidelities can mitigate this data quality/quantity-performance trade-off. Significantly lower average MAEs is achieved across all high-fi computed and the experimental predictions (Fig. 2a). The 5-fi MFGNet models, i.e., the models fitted using all available data, substantially improve on the MAE on the high-fi predictions over the 2-fi models, at the expense of a small increase in the MAE of the low-fi PBE predictions. Other 2-fi, 3-fi and 4-fi models, with and without PBE, were also explored (Table S1). The multi-fi models without PBE generally have higher MAEs than the multi-fi models with PBE, though they typically still outperform the 1-fi models. The 4-fi models that exclude the very small SCAN dataset outperform the 5-fi models across all non-SCAN fidelities, which indicates that the poor quality of the SCAN dataset may have degraded performance. The reduction in average MAE of the 4-fi models
over the 1-fi models range from $\sim 22\%$ for the experimental band gap to $\sim 45\%$ for the HSE band gap. Further, an increase in the number of fidelities also tends to improve model consistency, i.e., lower the standard deviation in the MAE.

The multi-fi MFGNet models significantly outperform prior ML models in the literature. The best reported GLLB-SC model has a RMSE of 0.95 eV, much higher than the 4-fi RMSE of 0.68 eV on the GLLB-SC predictions. For experimental band gaps, Zhuo et al. reported MAE and RMSE of 0.75 eV and 1.46 eV for a test set of 10 compounds using a support vector regression model; the MAE and RMSE for the 4-fi model on the experimental band gap of these compounds are 0.65 eV and 1.39 eV, respectively. Zhuo et al. also reported an RMSE of 0.45 eV on the entire experimental dataset, but the dataset contains duplicated band gaps for the same composition and thus is an inaccurate metric of model performance. We have also constructed a baseline 1-fi-stacked model, where a linear model is fitted for each high-fi dataset to the optimized 1-fi PBE model. This is akin to a model stacking approach and can be justified based on the relatively strong correlation between the high fidelity computed and PBE band gaps (Figure S1).

Multi-fi MFGNet models outperform the 1-fi-PBE+linear model, with especially large reductions in average MAEs of up to 38\% on arguably the most valuable experimental band gap predictions and 44-56\% on the GLLB-SC and HSE predictions. These results indicate that MFGNet framework is far better able to capture complex relationships between datasets of different fidelities.

To gain insights into the effect of low-fi and high-fi data size on model accuracy, 2-fi MFGNet models were developed using sampled subsets of each high-fi computed/experimental dataset ($N_{\text{high-fi}}$) together with different quantities of data from the low-fi PBE dataset ($N_{\text{PBE}}$). From Figure 2b-e, it may be observed that adding low-fi PBE data results in a significant decrease in the average MAEs of the high-fi predictions in all cases. The average MAEs of the 2-fi models follow an approximately linear relationship with $\log_{10} N_{\text{PBE}}$. With the exception of the SCAN 2-fi models, the magnitude of the slope decreases monotonically with an increase in $N_{\text{high-fi}}$, i.e., the largest improvements are observed in the most data-
constrained models. The nearly constant slope for the 2-fi SCAN models may be attributed to the extremely small size of the SCAN dataset as well as its strong correlation to the PBE dataset (Figure S1).

We compared the latent structural features extracted from the 1-fi and the 2-fi-exp models trained using 100 experimental data points without and with PBE data, respectively. The t-distributed Stochastic Neighbor Embedding (t-SNE)\textsuperscript{30} 2D projections of the latent structure features (Fig. 3a and b) show that the inclusion of the large PBE dataset in the 2-fi model results in superior structural representations that clearly separate structures with large band gap differences.

This separation can be further quantified by plotting the band gap difference $\Delta E_g$ against the distance in the normalized structural features $d_F$ between all 3,651,753 unique pairs of crystals in the experimental data, as shown in Fig. 3c and d). The 1-fi model for experimental band gaps has poor resolution, especially for large $\Delta E_g$. A wide $d_F$ range from 0.25 to 1 corresponds to $\Delta E_g \sim 10$ eV, and the correspondence between $d_F$ and $\Delta E_g$ is extremely noisy at low values. In contrast, the 2-fi-exp model exhibits an almost linear correspondence between $d_F$ and $\Delta E_g$ across the entire range of band gaps. Our conclusion is therefore that the inclusion of a large quantity of low-fidelity PBE data greatly assists in the learning of better latent structural features, which leads to substantially improved high-fidelity predictions. It should be noted, however, that a prerequisite for such improvements is that the low-fidelity data is of sufficient size and quality. For example, the 2-fi models without PBE perform worse than the 2-fi models with PBE (Table S1).

The MFGNet architecture also provides a natural framework to address another major gap in the computational materials property predictions - disordered crystals. The majority of known crystals exhibit site disorder. For example, of the $\sim 200,000$ crystals reported in the Inorganic Crystal Structure Database (ICSD), more than 120,000 are disordered crystals. Typically, the properties of disordered crystals are estimated by sampling low energy structures among a combinatorial enumeration of distinct orderings, usually within a supercell.
Figure 2: Mean absolute errors (MAEs) of MFGNet model predictions on ordered crystal band gaps. 

a, Performance of MFGNet models with different fidelity combinations. The 1-fi models were trained with each fidelity dataset only. The 2-fi models were trained using the PBE dataset with each higher fidelity dataset. The 4-fi models have the lowest average MAE on the experimental band gaps and were trained using the PBE, GLLB-SC, HSE and experimental datasets. The 5-fi models were trained using all available datasets. 

b, GLLB-SC, c, SCAN, d, HSE, e, experimental band gaps of 2-fi models trained using sampled datasets for each high-fidelity data and PBE data. The x-axis is plotted on a log scale and the shaded areas indicate one standard deviation of the MAE. $s$ indicates the slope for a linear fit of MAE to $\log_{10} N_{\text{PBE}}$. 
Figure 3: **Effect of including low-fidelity PBE data on latent structural features.** Two dimensional t-distributed Stochastic Neighbor Embedding (complexity = 1000) projection of features for **a**, 1-fi and **b**, 2-fi-PBE models trained using 100 experimental data points. The entire PBE dataset was used to train the 2-fi-PBE model. The markers are colored according to the experimental band gap. Plots of the experimental band gap difference ($\Delta E_g$) against normalized latent structural feature distance ($d_F$) in arbitrary units (a.u.) for the **c**, 1-fi and **d**, 2-fi-PBE models trained on all available experimental data. The dashed lines indicate the envelope of the maximum $\Delta E_g$ at each $d_F$. 
In MFGNet, we can use the learned length-16 elemental embeddings $W_Z$ directly as the node features. In such a scheme, disordered sites can be represented as a linear combination of the elemental embeddings as $W_{\text{disordered}} = \sum_{i=1}^{\text{species}} x_i W_{Z_i}$, where $x_i$ is the occupancy of species $i$ in the site and $W_{Z_i}$ is the element embedding for atomic number $Z_i$. Using the 4-fi model for the ordered crystals without further retraining, the MAE of the MFGNet predicted band gaps of the 278 disordered crystals in our experimental dataset is a respectable $0.63 \pm 0.14$ eV, similar to the MAE of the 1-fi-stacked model on ordered crystals. By retraining with the disordered experimental band gap dataset, the average MAE on the disordered band gaps decreases to $0.51 \pm 0.11$ eV, while that of the ordered crystals remains approximately the same ($0.37 \pm 0.02$ eV). The average MAEs for a retrained 1-fi MFGNet model on the disordered and ordered crystals are $0.55 \pm 0.13$ eV and $0.50 \pm 0.07$ eV, respectively. Clearly, the multi-fi approach improves on the performance on disordered crystals as well as ordered crystals.

To demonstrate the power of the disordered MFGNet models, band gap engineering data was extracted from the literature for $\text{Al}_x \text{Ga}_{1-x} \text{N}$, $\text{Cd}_{1-x} \text{Zn}_x \text{Se}$, $\text{Zn}_{1-x} \text{Mg}_x \text{O}$, and $\text{Lu}_3(\text{Ga}_x \text{Al}_{1-x})_5 \text{O}_{12}$. The band gaps of $\text{Lu}_3(\text{Ga}_x \text{Al}_{1-x})_5 \text{O}_{12}$ were not present and only the band gaps of the stoichiometric endpoints for the other systems were present in our experimental dataset. The 4-fi MFGNet model performs remarkably well, reproducing qualitative trends in all instances and achieving near quantitative accuracy for most systems. The 4-fi MFGNet model reproduces the concave relationship between $x$ and the change in band gap $\Delta E_g$ for $\text{Lu}_3(\text{Ga}_x \text{Al}_{1-x})_5 \text{O}_{12}$ (Figure 4d) reported by Fasoli et al. For $\text{Zn}_{1-x} \text{Mg}_x \text{O}$, a more pronounced concave relationship is predicted by the 4-fi MFGNet model compared to the experimental measurements. The band gap of ZnO is notoriously poorly estimated by DFT techniques, and even experimental measurements range from 3.1 to 3.4 eV across publications. An additional proof of concept for $\text{Ba}_y \text{Sr}_{1-y} \text{Co}_x \text{Fe}_{1-x} \text{O}_3 - \delta$ perovskite, a highly promising catalyst for the oxygen reduction reaction that exhibits disorder on multiple sites, is given in Figure S2.
Figure 4: Performance of Disordered MFGNet models. MFGNet predicted and experimental band gaps ($E_g$) as a function of composition variable $x$ in a, Al$_x$Ga$_{1-x}$N, b, Cd$_{1-x}$Zn$_x$Se, and c, Zn$_{1-x}$Mg$_x$O. d, Comparison of the change in band gap with respect to Lu$_3$Al$_5$O$_{12}$ ($\Delta E_g$) with $x$ in Lu$_3$(Ga$_x$Al$_{1-x}$)$_5$O$_{12}$. 
Data quality and quantity constraints are major bottlenecks to materials design. Multi-fidelity graph networks enable the efficient learning of latent structural features using large low-fidelity computed datasets to achieve vastly improved predictions from more costly computational methods and experiments. While crystal band gaps have been selected as the model problem in this work, the MFGNet framework is universal and readily applicable to other properties and to molecules. Two examples are provided in Figure S3, where a large number of low-fidelity molecule calculations are shown to lead to vast improvements in the high-fidelity energy predictions. Further, MFGNet’s ability to predict properties for disordered crystals opens up a vast space for *in silico* materials design that is extremely difficult or impossible to treat with existing *ab initio* computations or ML techniques.

**Methods**

**Multi-fidelity materials graph networks**

In materials graph networks, atoms and bonds are represented as nodes and edges in an undirected graph as \((V, E, u)\). The atom attributes \(V\) are the atomic numbers \(Z \in \mathbb{N}\). Each atom attribute is associated with a row vector \(W_{Z_i} \in \mathbb{R}^{16}\) in the element embedding matrix \(W_Z = [W_0; W_1; W_2; \ldots; W_{94}]\) where \(W_0\) is a dummy vector. The bond attribute is the set of Gaussian-expanded distances. For the \(k\)-th bond in the structure, the attributes are

\[
e_{k,m} = \exp\left(-\frac{(d_k - \mu_m)^2}{\sigma^2}\right), \forall d_k \leq R_c
\]

where \(d_k\) is the length of the bond \(k\) formed by atom indices \(r_k\) and \(s_k\), \(R_c\) is the cutoff radius and \(\mu_m = \frac{m}{n_{bf} - 1}\mu_{max}\) for \(m = \{0, 1, 2, \ldots, n_{bf} - 1\}\) and \(n_{bf}\) is the number of bond features. In this work, \(R_c = 5\,\text{Å}, \mu_{max} = 6\,\text{Å},\) and \(n_{bf} = 100\). The graphs are constructed using an edge list representation, and the edge set of the graph is represented as \(E = \{(e_k, r_k, s_k)\}\). The state attributes \(u\) are fidelity levels \(F \in \mathbb{N}\). Similar to atom attributes, fidelity \(F_i\) is associated
with a row vector $W^f_{F_i}$ in the fidelity embedding matrix $W_F = [W^f_0, W^f_1, W^f_2, W^f_3, W^f_4]$. Both embedding matrices $W_Z$ and $W_F$ are trainable in the models, except in disordered models where the elemental embedding matrix $W_Z$ is fixed to previously obtained values.

In each graph convolution layer, the graph networks are propagated sequentially as follows:

1. The attributes of each bond $k$ in the graph are updated as

$$e'_k = \phi_e(v_{s_k} \oplus v_{r_k} \oplus e_k \oplus u)$$

where $\phi_e$ is the bond update function, $v_{s_k}$ and $v_{r_k}$ are the atomic attributes of the two atoms forming the bond $k$, and $\oplus$ is the concatenation function.

2. Each atom $i$ is then updated as

$$v'_i = \phi_v(\bar{v}_i^e \oplus v_i \oplus u)$$

where $\phi_v$ is the atomic update function, and $\bar{v}_i^e = \text{average}_k(e'_k), \forall r_k = i$ is the averaged bond attributes from all bonds connected to atom $i$.

3. Finally, the state attributes are updated as

$$u' = \phi_u(\bar{u}^e \oplus \bar{u}^v \oplus u)$$

where $\phi_u$ is the state update function, and $\bar{u}^e = \text{average}_k(e'_k)$ and $\bar{u}^v = \text{average}_i(v'_i)$ are the averaged attributes from all atoms and bonds, respectively.

The update functions $\phi_e$, $\phi_v$ and $\phi_u$ are multi-layer perceptron models with $[64, 32, 32]$ hidden neurons. In this work, three graph network layers are stacked to increase the models’ predictive power.
Data Collection and Processing

The PBE dataset comprising 52,348 crystal structures with band gaps were obtained from Materials Project on Jun 1 2019 using the Materials Application Programming Interface in the Python Materials Genomics (pymatgen) library. The GLLB-SC band gaps from Castelli et al. were obtained via MPContribs. The total number of GLLB-SC band gaps is 2,290 after filtering out materials that do not have structures in the current Materials Project database and those that failed the graph computations due to abnormally long bond (>5 Å). The GLLB-SC data all have positive band gaps due to the constraints applied in the structure selection in the previous work. The SCAN band gaps for 472 nonmagnetic materials were obtained from Borlido et al. The HSE band gaps with corresponding Materials Project structures were downloaded from the MaterialGo website. After filtering out ill-converged calculations and those that have a much smaller HSE band gap compared to the PBE band gaps, 6,030 data points remain, of which 2,775 are metallic. Finally, the experimental band gaps were obtained from the work by Zhuo et al. As this data set only contains compositions, the experimental crystal structure for each composition was obtained by looking up the lowest energy polymorph for a given formula in the Materials Project, followed by cross-referencing with the corresponding Inorganic Crystal Structure Database (ICSD) entry. Further, as multiple band gap can be reported for the same composition in this data set, the band gaps for the duplicated entries were averaged. In total, 2,703 ordered (938 binary, 1306 ternary and 459 quaternary) and 278 disordered (41 binary, 132 ternary and 105 quaternary) structure-band gap pairs were obtained.

Model training

We split the data into 80%-10%-10% train-validation-test ratios randomly and repeated the splitting at least five times. The models are trained on the training data for a maximum of 1,500 epochs and stop early when the validation error does not reduce for consecutive 500 epochs. The final model performances were evaluated on the test set and reported in
this work. Three graph convolution layers are used in the model training, with $[64, 64, 32]$ hidden units in each layer. A learning rate of $10^{-3}$ is used with Adam optimizer$^{12}$ and the batch size for the training is set to 128.

Acknowledgement

The authors acknowledge the support from the Materials Project, funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division under contract no. DE-AC02-05-CH11231: Materials Project program KC23MP. C.C. thanks Dr. Matthew Horton for his assistance with the GLLB-SC data set.

Author contributions C.C. and S.P.O. conceived the idea and designed the work. C.C. implemented the models and performed the analysis. S.P.O. supervised the project. Y.Z., W.Y. and X.L. helped with the data collection and analysis. C.C. and S.P.O. wrote the manuscript. All authors contributed to the discussion and revision.

Competing Interests The authors declare that they have no competing financial interests.

Correspondence Correspondence and requests for materials should be addressed to S.P.O. (email: ongsp@eng.ucsd.edu).

Data availability

All data generated and analysed during the current study are available from the corresponding authors on reasonable request.

References

(1) Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine Learning for Molecular and Materials Science. *Nature* **2018**, *559*, 547–555.
(2) Chen, C.; Zuo, Y.; Ye, W.; Li, X.; Deng, Z.; Ong, S. P. A Critical Review of Machine Learning of Energy Materials. *Advanced Energy Materials* **2020**, 10, 1903242.

(3) Battaglia, P. W. et al. Relational Inductive Biases, Deep Learning, and Graph Networks. *arXiv:1806.01261 [cs, stat] 2018*.

(4) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chemistry of Materials* **2019**, 31, 3564–3572.

(5) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, 77, 3865–3868.

(6) Chevrier, V. L.; Ong, S. P.; Armiento, R.; Chan, M. K. Y.; Ceder, G. Hybrid Density Functional Calculations of Redox Potentials and Formation Energies of Transition Metal Compounds. *Physical Review B* **2010**, 82, 075122.

(7) Heyd, J.; Scuseria, G. E. Efficient Hybrid Density Functional Calculations in Solids: Assessment of the Heyd–Scuseria–Ernzerhof Screened Coulomb Hybrid Functional. *The Journal of Chemical Physics* **2004**, 121, 1187–1192.

(8) Zhang, Y.; Kitchaev, D. A.; Yang, J.; Chen, T.; Dacek, S. T.; Sarmiento-Pérez, R. A.; Marques, M. A. L.; Peng, H.; Ceder, G.; Perdew, J. P.; Sun, J. Efficient First-Principles Prediction of Solid Stability: Towards Chemical Accuracy. *npj Computational Materials* **2018**, 4, 9.

(9) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Physical Review Letters* **2018**, 120, 145301.

(10) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Physical Review Letters* **2007**, 98, 146401.
(11) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Physical Review Letters* **2010**, *104*, 136403.

(12) Zuo, Y.; Chen, C.; Li, X.; Deng, Z.; Chen, Y.; Behler, J.; Csányi, G.; Shapeev, A. V.; Thompson, A. P.; Wood, M. A.; Ong, S. P. Performance and Cost Assessment of Machine Learning Interatomic Potentials. *The Journal of Physical Chemistry A* **2020**, *124*, 731–745.

(13) 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 Materials* **2013**, *1*, 011002.

(14) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): Assessing the Accuracy of DFT Formation Energies. *npj Computational Materials* **2015**, *1*, 15010.

(15) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid Functionals Based on a Screened Coulomb Potential. *The Journal of Chemical Physics* **2003**, *118*, 8207–8215.

(16) Hachmann, J.; Olivares-Amaya, R.; Atahan-Evrenk, S.; Amador-Bedolla, C.; Sánchez-Carrera, R. S.; Gold-Parker, A.; Vogt, L.; Brockway, A. M.; Aspuru-Guzik, A. The Harvard Clean Energy Project: Large-Scale Computational Screening and Design of Organic Photovoltaics on the World Community Grid. *The Journal of Physical Chemistry Letters* **2011**, *2*, 2241–2251.

(17) Hellwege, K. H.; Green, L. C. Landolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology. *American Journal of Physics* **1967**, *35*, 291–292.

(18) Pilania, G.; Gubernatis, J. E.; Lookman, T. Multi-Fidelity Machine Learning Models
for Accurate Bandgap Predictions of Solids. *Computational Materials Science* 2017, *129*, 156–163.

(19) Batra, R.; Pilania, G.; Uberuaga, B. P.; Ramprasad, R. Multifidelity Information Fusion with Machine Learning: A Case Study of Dopant Formation Energies in Hafnia. *ACS Applied Materials & Interfaces* 2019, *11*, 24906–24918.

(20) Gritsenko, O.; van Leeuwen, R.; van Lenthe, E.; Baerends, E. J. Self-Consistent Approximation to the Kohn-Sham Exchange Potential. *Physical Review A* 1995, *51*, 1944–1954.

(21) Kuisma, M.; Ojanen, J.; Enkovaara, J.; Rantala, T. T. Kohn-Sham Potential with Discontinuity for Band Gap Materials. *Physical Review B* 2010, *82*, 115106.

(22) Castelli, I. E.; Hüser, F.; Pandey, M.; Li, H.; Thygesen, K. S.; Seger, B.; Jain, A.; Persson, K. A.; Ceder, G.; Jacobsen, K. W. New Light-Harvesting Materials Using Accurate and Efficient Bandgap Calculations. *Advanced Energy Materials* 2015, *5*, 1400915.

(23) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Physical Review Letters* 2015, *115*, 036402.

(24) Borlido, P.; Aull, T.; Huran, A. W.; Tran, F.; Marques, M. A. L.; Botti, S. Large-Scale Benchmark of Exchange–Correlation Functionals for the Determination of Electronic Band Gaps of Solids. *Journal of Chemical Theory and Computation* 2019, *15*, 5069–5079.

(25) Jie, J.; Weng, M.; Li, S.; Chen, D.; Li, S.; Xiao, W.; Zheng, J.; Pan, F.; Wang, L. A New MaterialGo Database and Its Comparison with Other High-Throughput Electronic Structure Databases for Their Predicted Energy Band Gaps. *Science China Technological Sciences* 2019,
(26) Zhuo, Y.; Mansouri Tehrani, A.; Brgoch, J. Predicting the Band Gaps of Inorganic Solids by Machine Learning. The Journal of Physical Chemistry Letters 2018, 9, 1668–1673.

(27) Perdew, J. P.; Levy, M. Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities. Physical Review Letters 1983, 51, 1884–1887.

(28) Davies, D. W.; Butler, K. T.; Walsh, A. Data-Driven Discovery of Photoactive Quaternary Oxides Using First-Principles Machine Learning. Chemistry of Materials 2019, 31, 7221–7230.

(29) Morales-García, Á.; Valero, R.; Illas, F. An Empirical, yet Practical Way To Predict the Band Gap in Solids by Using Density Functional Band Structure Calculations. The Journal of Physical Chemistry C 2017, 121, 18862–18866.

(30) van der Maaten, L.; Hinton, G. Visualizing Data Using T-SNE. Journal of Machine Learning Research 2008, 9, 2579–2605.

(31) Chen, H.; Chen, K.; Drabold, D. A.; Kordesch, M. E. Band Gap Engineering in Amorphous AlxGa1-xN: Experiment and Ab Initio Calculations. Applied Physics Letters 2000, 77, 1117–1119.

(32) Santhosh, T. C. M.; Bangera, K. V.; Shivakumar, G. K. Band Gap Engineering of Mixed Cd(1-x)Zn (x) Se Thin Films. Journal of Alloys and Compounds 2017, 703, 40–44.

(33) Rana, N.; Chand, S.; Gathania, A. K. Band Gap Engineering of ZnO by Doping with Mg. Physica Scripta 2015, 90, 085502.

(34) Fasoli, M.; Vedda, A.; Nikl, M.; Jiang, C.; Uberuaga, B. P.; Andersson, D. A.; McClellan, K. J.; Stanek, C. R. Band-Gap Engineering for Removing Shallow Traps in
Rare-Earth Lu$\{\cdot3\}Al$\{\cdot5\}O$\{\cdot12\}$ Garnet Scintillators Using Ga$\{\cdot3+\}$ Doping. Physical Review B 2011, 84, 081102.

(35) Harun, K.; Salleh, N. A.; Deghfel, B.; Yaakob, M. K.; Mohamad, A. A. DFT + U Calculations for Electronic, Structural, and Optical Properties of ZnO Wurtzite Structure: A Review. Results in Physics 2020, 16, 102829.

(36) Kamarulzaman, N.; Kasim, M. F.; Chayed, N. F. Elucidation of the Highest Valence Band and Lowest Conduction Band Shifts Using XPS for ZnO and Zn0.99Cu0.01O Band Gap Changes. Results in Physics 2016, 6, 217–230.

(37) Shao, Z.; Haile, S. M. A High-Performance Cathode for the next Generation of Solid-Oxide Fuel Cells. Nature 2004, 431, 170–173.

(38) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. Computational Materials Science 2013, 68, 314–319.

(39) Ong, S. P.; Cholia, S.; Jain, A.; Brafman, M.; Gunter, D.; Ceder, G.; Persson, K. A. The Materials Application Programming Interface (API): A Simple, Flexible and Efficient API for Materials Data Based on REpresentational State Transfer (REST) Principles. Computational Materials Science 2015, 97, 209–215.

(40) Huck, P.; Jain, A.; Gunter, D.; Winston, D.; Persson, K. A Community Contribution Framework for Sharing Materials Data with Materials Project. 2015 IEEE 11th International Conference on E-Science. 2015; pp 535–541.

(41) Hellenbrandt, M. The Inorganic Crystal Structure Database (ICSD)—Present and Future. Crystallography Reviews 2004, 10, 17–22.
(42) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980* 2014,