Neural Computing and Applications (2022) 34:4625–4632
https://doi.org/10.1007/s00521-021-06616-0

ORIGINAL ARTICLE

Graph neural network for Hamiltonian-based material property prediction

Hexin Bai[1] · Peng Chu[2] · Jeng-Yuan Tsai[1] · Nathan Wilson[3] · Xiaofeng Qian[3] · Qimin Yan[1] · Haibin Ling[4]

Received: 2 May 2021 / Accepted: 4 October 2021 / Published online: 13 November 2021

Abstract

Development of next-generation electronic devices calls for the discovery of quantum materials hosting novel electronic, magnetic, and topological properties. Traditional electronic structure methods require expensive computation time and memory consumption, thus a fast and accurate prediction model is desired with increasing importance. Representing the interactions among atomic orbitals in material, a Hamiltonian matrix provides all the essential elements that control the structure–property correlations in inorganic compounds. Learning of Hamiltonian by machine learning therefore offers an approach to accelerate the discovery and design of quantum materials. With this motivation, we present and compare several different graph convolution networks that are able to predict the band gap for inorganic materials. The models are developed to incorporate two different features: the information of each orbital itself and the interaction between each other. The information of each orbital includes the name, relative coordinates with respect to the center of super cell and the atom number. The interaction between orbitals is represented by the Hamiltonian matrix. The results show that our model can get a promising prediction accuracy with cross-validation.

Keywords Graph neural network · Graph convolution network · Physics · Material

1 Introduction

It has long been the interest of physicist to discover the electronic, magnetic, and topological properties based on fundamental physics and electronic structure. Among all the properties, one specifically intrigues us in this paper is the band gap which directly reveals the conductivity of the material. On the other hand, the Hamiltonian matrix calculated by first-principles in tight-binding setup could be used to represent fundamental physics and electronic structure [36]. Traditionally, it costs days or even weeks to compute the band gap in first-principles calculation using Hamiltonian matrix. Instead, we incorporate Hamiltonian matrix of inorganic crystalline systems in our network-based machine learning framework for band gap prediction so that band gap could be predicted in a much faster and easier way.

Traditionally, convolutional neural network (CNN) [21] has been used in real-space wavefunction-based analysis of small molecules and orthogonal systems. However, for nonorthogonal grids and large molecules, CNN-based

✉ Hexin Bai
tug76276@temple.edu

Peng Chu
pengchu@microsoft.com

Jeng-Yuan Tsai
jeng-yuan.tsai@temple.edu

Nathan Wilson
wilsonnater@tamu.edu

Xiaofeng Qian
feng@tamu.edu

Qimin Yan
qiminyan@temple.edu

Haibin Ling
haibin.ling@stonybrook.edu

[1] Temple University, Philadelphia, USA

[2] Microsoft, Redmond, USA

[3] Texas A&M University, College Station, USA

[4] Stony Brook University, New York, USA

methods are insufficient for the proper modeling since the locality in infinite or nonorthogonal systems differs substantially from that in finite orthogonal systems. In other words, the great diversity in condensed morphology of materials limits the applications of current machine learning techniques such as CNN. More specifically, the dimension of the Hamiltonian matrix is highly variant. Thus, the Hamiltonian matrix does not fit directly in the fixed local neighborhood of CNN. And the variant dimension of Hamiltonian matrix prohibits generalization of models trained on one Hamiltonian matrix to other Hamiltonian matrix with different dimension.

Instead of treating Hamiltonian matrix as an image in traditional CNN, we use graph to represent it. The reason is that basis of atoms and mutual interactions in quantum materials naturally encourage using the general graph representation for materials. In particular, locality in quantum material systems can be defined as graph nodes apart from their actual grid and scale in real space. And the interaction between elements represented by Hamiltonian matrix is graph edges. In such a way, we can easily deal with the variant dimension of number of elements and the Hamiltonian matrix built from it. This whole graph builds up the input of our machine learning pipeline, while the output is the band gap. We set up a threshold for band gap to make it a classification since whether the band gap is relatively large or small is more of interest for material study rather than a exact value.

With this input and output, we are able to leverage the recent emerging graph convolutional networks (GCN) or graph neural network (GNN) [4, 16, 19, 31]. In this paper, we explore two different graph convolutional network architectures to classify the band gap of the quantum materials using their Hamiltonian matrices and atom basis. The message passing graph network [11] conducts the information aggregation of neighboring nodes in the graph. On the other hand, Chebyshev convolution [7] leverages the Chebyshev polynomial to accelerate the convolutional operation in the spatial domain. The two methods are tested on our collected dataset to predict the band gap. Detailed procedure is plotted in Fig. 1. Compared with traditional hand-crafted feature-based methods, the GCN-based ones have better performance on this binary classification task.

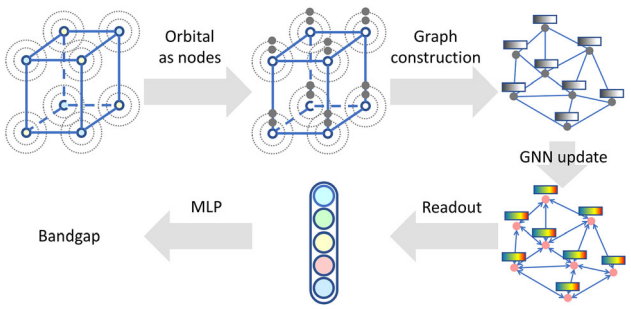

Fig. 1 Visualization for pipeline

2 Related work

Machine learning techniques provide a novel opportunity to speed up materials discovery by utilizing data-driven paradigms [10, 33, 37]. Instead of numerically solving complex systems with quantum interactions, physical quantities are statistically estimated based on a reference set of known solutions. Machine learning, especially supervised learning, has been applied to predictions of phase stability [26], crystal structure [6, 15], electronic structure [23, 33], molecule atomization energies [14], effective potential for molecule dynamics [2, 3], and energy functional for density functional theory [40].

More specifically, for the band gap prediction by machine learning, deep neural networks have been used [39, 43]. Besides, Joohwi Lee et al. [22] used linear regression and support vector machine to predict band gap. Similarly, support vector machine was employed by Ya Zhuo et al. [45]. Ghanshyam Pilania et al. [35] leveraged Gaussian process regression. Rajan et al. [38] used many algorithms including kernel ridge regression, support vector machine, Gaussian process regression and bagging. Similar to the previous works, Olsthoorn et al. [30] used kernel ridge regression and deep neural network for band gap prediction. Pilania et al. [34] later published another paper using the popular kernel ridge regression. The tree-based method was also used: Logan Ward et al. [1] used random forest to predict band gap. But all of the work mentioned above ignore the Hamiltonian matrix which is important in deciding the band gap of material. One reason would be the variant dimension of it. In fact, some authors mentioned above intentionally limit the scope of the material they study to fix the dimension of the elements. These all imply that a graph-based learning which incorporated the variant dimension of Hamiltonian matrix is the choice.

The variant dimension of Hamiltonian matrix makes it ideal for GCN. Meanwhile, GCN has great advance recently in pattern recognition and data mining. Generally speaking, there are mainly two structures for supervised learning in GCN: spectral method and spatial method. For spectral method, two representative methods are [19] and [7]. They tried to leverage the eigenspace of the graph Laplacian matrix to make the prediction. Such a method has a more profound mathematical theory foundation but is bad at transferring from the graph learned to the graph unseen in the test set. For spatial methods, Justin Gilmer et al. laid the foundation in their recent work [11]. In the same year, William Hamilton et al. published [13]. Both of

these methods focus on the neighborhood of node in the graph and hence gain the ability to transfer from the graph learned to the graph unseen. Later on, spatial method becomes the mainstream of the study in the field of GCN with many works contributed. For a more comprehensive literature review, Wu gave a summary [41]. In our paper, we used the two most representative GCN: [7] from spectral method and [11] from spatial method to test their difference in performance.

3 Method

3.1 Problem formulation

In materials science, the material's band gap is an important property governing whether the material is metal or non-metal. In this study, we aim to use GCN to predict the band gap given the Hamiltonian of the material. Band gap is described by a nonnegative real number, $E_g \in \mathbb{R}$ and $E_g \geq 0$. To simplify the problem, threshold is applied to split E_g into two categories. Hence, we have $c = 1$ for $E_g > 0.2$ and $c = 0$ for $E_g \leq 0.2$. The band gap is set to 0.2 because it creates a balanced binary label for our dataset. Namely, it splits the dataset in half. These two classes represent the metal and non-metal classes as the learning target. Finally, the learning problem is defined with a cross-entropy loss:

$$\hat{\theta} = \text{argmin}_\theta - \big(f_\theta(x)\log(c) + (1 - f_\theta(x))\log(1-c)\big) \tag{1}$$

where $\hat{c} = f_{\hat{\theta}}(x)$ is the prediction, x is the input representation of Hamiltonian, and $f_\theta(\cdot)$ is the function with trainable parameters θ.

3.2 Hamiltonian matrix

We will focus on the 2D Hamiltonian matrix that represents the fundamental physics of materials. Specifically, Hamiltonian matrix of a physical system contains all the operators of the kinetic and potential energies. Let the Hamiltonian operator $\hat{H}$ for an N-particle condensed matter system be

$$\hat{H} = \sum_{n=1}^{N} \hat{T}_n + \hat{V}, \tag{2}$$

where $\hat{T}_n$ indicates the kinetic energy operators for each particle, and $\hat{V}$ is the potential energy operator between particles.

To facilitate the calculation, the Hamiltonian operator can be represented as a 2D numerical matrix. In detail, the representation of an operator can be obtained through the integral with the basis of a Hilbert space. In the condense matter system, the wavefunctions $\{\varphi_i\}$ of orbitals from all atoms in the material system can form a Hilbert space. Therefore, the element in the matrix representation of Hamiltonian H_{ij} can be calculated from

$$H_{ij} = \int \varphi_i \hat{H} \varphi_j d^3 r, \tag{3}$$

resulting in an $M \times M$ Hamiltonian matrix $H = \{H_{ij}\}$ with M as the total number of orbitals. In this paper, the Hamiltonian is computed from a super cell of $7 \times 7 \times 7 = 343$ unit cells. Each cell contains several atoms with a finite number of orbitals. In Fig. 2, we visualize the Hamiltonian matrix for a sample in the Li-Al-Si material system with $M = 243$. Note that, in order to keep consistency with our experiment, we consider only the center $3 \times 3 \times 3 = 27$ of the 343 for computation cost. In the example of Fig. 2, with 9 orbitals, we obtained a square matrix of dimension $27 \times 9 = 243$ filled with the real part value. The imaginary parts are so small that we can neglect them. The sidebar of Fig. 2 ranges from 0.0 to 1.4, while it represents the energies of the orbitals in material.

3.3 Graph representation of Hamiltonian

Hamiltonian contains more concentrated information and lower input dimension than wavefunctions or charge density. Thus, directly using Hamiltonian for prediction may reduce the model complexity and relieve the demand of big data for model training. However, due to different atom composition in each condensed matter system, the size of Hamiltonian varies greatly. To handle the diversity of input

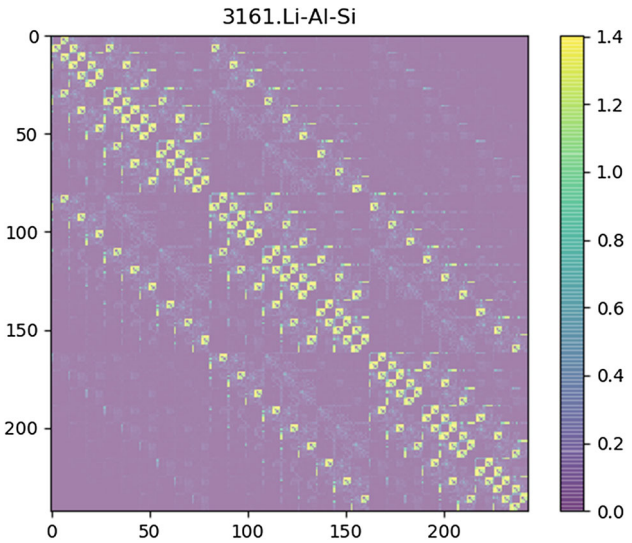

Fig. 2 The Hamiltonian matrix of Li-Al-Si. The side bar from 0.0 to 1.4 represents the energies of orbitals in material

dimension, we propose using a graph to store the Hamiltonian matrix in this learning task.

A weighted undirected graph is constructed from the Hamiltonian matrix to encode interaction and symmetric information of the quantum system. As indicated in Eq. 3, the matrix element H_{ij} represents the interaction intensity between the i-th orbital and the j-th orbital in the condensed matter system. Then, if the orbitals are represented by M vertices in a graph, such as $\mathcal{V} = \{\mathbf{v}_i\}_{i=1:M}$, naturally, the interaction between orbitals can be described by the edges in the graph. Noting that, according to Eq. 3, interactions exist for each pair of orbitals in the system, which results in a complete graph. Here, we only consider the interaction stronger than a threshold τ_h to reduce the computational burden. Finally, a graph $\mathcal{G} = (c, \mathcal{V}, \mathcal{E})$, with $\mathcal{E} = \{(\mathbf{e}_{ij}, i, j)\}_{i,j\in 1,2,\ldots M}$ for M vertices, is built for each Hamiltonian matrix. Specifically, $\mathbf{e}_{ij}$ is the edge which represents the interaction between i-th and j-th orbitals; its weight is the real part of complex number H_{ij} in Hamiltonian matrix. In practice, the scale of the imaginary part is negligible comparing with the real part. Hence, the real part of the complex number is a close approximation to the modulus of the complex number.

For each node inside the graph, we have the following node feature vector representation. In Table 1, a summary of feature components for the node and edges is listed. Two types of features are included here: the node feature and edge feature. For the three node features, they are concatenated into vectors. The edge feature is organized as the graph adjacency matrix. The "Type" here illustrates the data type of each feature, such as vector or scalar, integer or real number. The "Description" here represents the physics meaning of the features. The collection of node features for all nodes in $3 \times 3 \times 3$ unit cells of a Li-Al-Si sample is visualized in Fig. 3. The sidebar corresponds to the three features of atomic number, atom coordinate, and orbit type. More specifically, the one-hot vector of orbit type ranges from 0 to 1; atom number ranges from 3 to 14; and coordinate ranges from -2 to 2. Combining them, we have a range from -2 to 14.

With the graph represented Hamiltonian, we investigate two types of GCNs to learn the information for the band gap prediction.

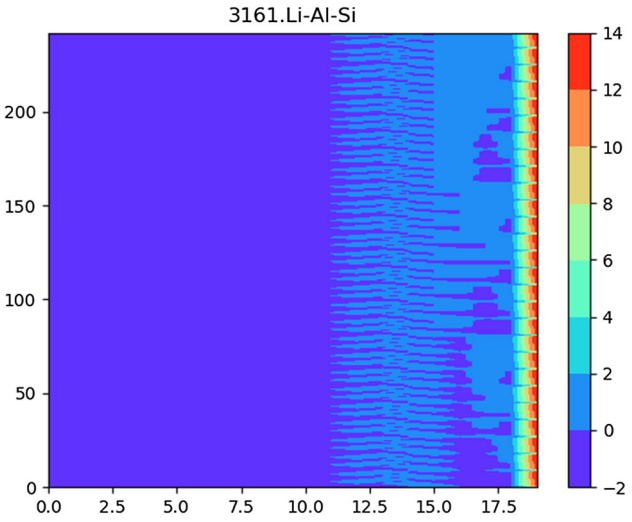

Fig. 3 Visualization of node features for Li-Al-Si. Each row of the matrix represents a node feature in the $3 \times 3 \times 3$ unit cells. The sidebar from -2 to 14 incorporates the one-hot encoding of orbit type, atom number, and coordinates with detail in Table 1

3.4 Message passing graph network

A single graph convolutional block is composed of update functions ϕ and aggregation functions ρ, such as

$$\begin{aligned} \mathbf{x}_{ik} &= \psi(\mathbf{e}_{ik}, \mathbf{v}_k), \ k \in \mathcal{N}(i) \\ \bar{\mathbf{x}}_i &= \rho_{\text{local}}\Big(\bigcup_{k\in\mathcal{N}(i)} \mathbf{x}_{ik}\Big) \\ \bar{\mathbf{v}} &= \rho_{\text{global}}\Big(\bigcup_{i} \bar{\mathbf{x}}_i\Big) \\ \hat{c} &= \phi^c(\bar{\mathbf{v}}) \end{aligned} \tag{4}$$

where $\mathcal{N}(i)$ stands for the neighborhood of node i. The function ρ is order-invariant aggregation. ρ_{local} aggregates information from all the edges connecting to the node i. ρ_{global} summarizes the information globally. The function ϕ^c predicts the global attribute. The $\mathbf{x}_{ik}$, updated by function ψ, is a learned vector representation for each node that could be updated multiple times.

This general framework can be implemented with different flexible variants. In [11], the message passing neural network is proposed to allow long range interactions

Table 1 Feature component used in the GCN methods

	Physics feature	Data type	Physics meaning
Node	Atomic number	1 Integer	Number of protons
	Atom coordinates	3 Real	3D location in the unit cell
	Orbit type	1×16 One hot	$s, p, d, \ldots$
Edge	Interaction	1 Real	Real part of H_{ij}

between nodes in the graph for molecular properties prediction. A modified version is implemented here where both ψ and ϕ^c are multilayer perception; ρ_{local} is local average function; ρ_{global} is global average function with random dropout of nodes.

3.5 Chebyshev convolution

For k-th vertex-wised signal $\mathbf{v}_k$, the convolution operation on graph $\mathcal{G}$ such as $\mathbf{v}_k * \mathcal{G}$ can also be defined in the frequency domain. To analyze graph $\mathcal{G}$ in Fourier domain, one essential operation is to obtain the graph Laplacian $L = D - E$ where $E = \{\mathbf{e}_{ij}\}$ and D is the diagonal degree matrix with $D_{ii} = \sum_j \mathbf{e}_{ij}$ or as a normalized form $L = I_n - D^{(-1/2)}ED^{(-1/2)}$. The Laplacian can be diagonalized by its graph Fourier modes $U = [u_0, u_1, \ldots, u_{M-1}]$ such that $L = U\Lambda U^T$ where $\Lambda = \text{diag}\big([\lambda_0, \lambda_1, \ldots, \lambda_{M-1}]\big)$ are the frequencies of the graph. Finally, a signal $\mathbf{v}_k$ is filtered by g_θ as

$$\mathbf{y}_k = g_\theta(L)\mathbf{v}_k = g_\theta\big(U\Lambda U^T\big)\mathbf{v}_k = Ug_\theta(\Lambda)U^T\mathbf{v}_k. \tag{5}$$

Convolution with constant neighbors could use nonparametric filter such as $g_\theta(\Lambda) = diag(\theta)$ where $\theta \in \mathbb{R}^M$ is trainable variables. However, it is not localized in space. Therefore, polynomial parametrization is used to construct the localized filters $g_\theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k \Lambda^k$ which limit the shortest path distance to K (e.g., within the K-th order neighbors of a vertex) ([7]). In order to further accelerate the computation of the multiplication with the Fourier basis U, Chebyshev polynomial $T_k(x)$ is used to build the filter

$$g_\theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k T_k\big(\tilde{\Lambda}\big). \tag{6}$$

As a result, the filtering operation can be written as

$$\mathbf{y}_k = g_\theta(L)x = \sum_{k=0}^{K-1} \theta_k T_k\big(\tilde{L}\big)\mathbf{v}_k \tag{7}$$

where θ_k is the trainable parameters for a single output channel.

4 Experiments

4.1 Dataset

We collected 530 half-Heusler compounds from the Materials Project database [17] using the data mining approach. A total of 233 samples are used after cleaning the dataset. Each of the generated raw Hamiltonian sample contains three atoms where each atom has a maximum of 16 orbitals. The Hamiltonian matrices are all calculated within $7 \times 7 \times 7 = 343$ unit cells. The real values of target band gap fall in the range of $[0, 5.6]$. We choose a threshold 0.2 to produce binary labels ($E_g > 0.2$ as $c = 1$ and $E_g \leq 0.2$ as $c = 0$). This results in a balanced subset with 116 positive samples and 117 negative samples.

4.2 Features

We generate two different features for GCN-based methods and shallow methods, respectively. The feature used in GCN-based approaches is already presented before.

For shallow methods, a set of fixed length features is created to include both atomic and interaction information. To limit the total dimension of feature, only the Hamiltonian from the center unit cell is selected for use. Zeros padding is used to accommodate size variation in the Hamiltonian for different samples. In detail, all Hamiltonian are embedded in the square matrix of size 48×48 where each side corresponds to the three atoms each with 16 orbitals. The interaction features are the vectorization of the square matrix, which results in dimension of 2,304. The atomic features are the concatenation of the atomic number and atom coordinate feature in Table 1. Finally, the combination of the two features results in a set of features with a dimension of $3 + 3 \times 3 + 2304 = 2316$.

4.3 Experiment settings

Five popular shallow classification methods are evaluated, including decision tree, Naive Bayes, Multilayer perception, SVM and random forest. Three variations in the features are used in these experiments: (1) Interaction only, (2) Atom only, and (3) Both atom and interaction. Since the class distribution is balanced, the baseline performance of random output is 50%. Reported results are binary classification accuracy by fivefold cross-validation.

A two-layer MPNN and a three-layer Chebyshev convolutional network are built upon the Hamiltonian matrix. For Chebyshev convolutional network, the convolutional filter size in those three layers is chosen as {1, 2, 2}, respectively, to gradually increase the receptive field. Leaky-ReLU is adapted for nonlinear activation. At last, a global average pooling layer followed by a softmax layer outputs the probabilities of input graph that belong to the two categories. For both methods, Adam optimizer with a fixed learning rate of 0.001 and weight decay of 5e-4 is used. The training is performed up to 2,000 epochs. The two GCNs are implemented in PyTorch Geometric and trained on an Nvidia Titan X GPU.

4.4 Results

We report the classification accuracy of the shallow methods using the fivefold cross-validation in Table 2. For "Method," it illustrates the list of the methods we used. The "Interaction" column represents the result of considering only orbitals' interaction while leaving atoms' information apart. Conversely, the "Atom" column represents the result of considering only atoms' information (atom number, atom coordinate, etc.) while leaving the orbitals interaction apart. The meaning of "Atom + Interaction" in traditional machine learning methods represents the result of concatenating atoms and interactions as input feature vectors. As for the graph neural network, "Atom + Interaction" represents the result of using the whole graph with atoms as nodes and interactions as adjacency matrix for input. Among the different variations of features, using both atom and interaction usually achieves the best performance. The Interaction features alone achieve the similar performance with the combination of Interaction and Atom feature. This demonstrates the importance of using the Hamiltonian to represent the material. The experiments on shallow methods confirm that the interaction embedded in the Hamiltonian contains the key information for the band gap prediction. Therefore, using graph to model this interaction structure of the Hamiltonian and applying GCN for learning process provide an advanced solution to this task.

The learning curves of MPNN and Chebyshev network are shown in Fig. 4. The same cross-validation set generated with the same random seed is used here for fair comparison. Without the random dropout module, the Chebyshev network clearly has over fitting on the validation accuracy curve with nearly 68% accuracy while it has 98% accuracy on the training set. By applying the random dropout module, there is only 10% difference of accuracy in training and validation for MPNN. We also report numerical performance of all fivefold in Table 2.

We also explore the case that includes the neighboring unit cell for the classification task. When only center unit cell was used, one can observe from the results that the GCN-based method is better than traditional shallow methods. This is due to more accurate and compact graph representation of the Hamiltonian data. By further including the neighboring unit cells, feature dimension will dramatically increase to 20,000 for shallow methods. This high dimension input is unacceptable for shallow methods before effective dimension reduction technique is applied. But the GCN-based methods can handle this case naturally. As shown in Table 2, Chebyshev convolution-based network achieves similar performance with the eight times larger input. With the limited training samples, the high dimensional input may not directly benefit the prediction. When training with sufficient data, we can expect the GCN to achieve much better performance in learning the periodic structure information in the Hamiltonian for the physical properties prediction.

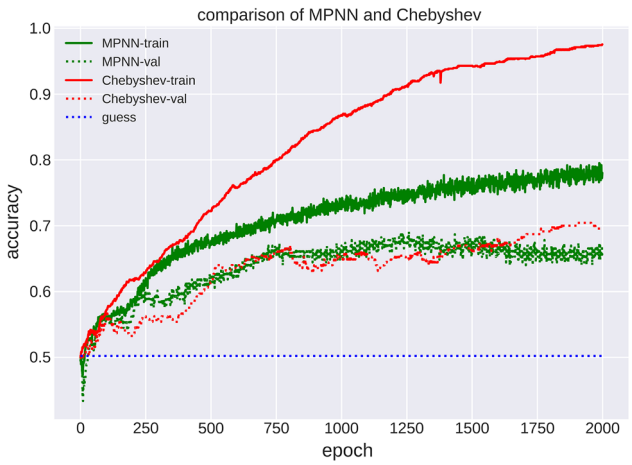

Fig. 4 Learning curves of the two GCN-based methods on onefold

5 Conclusion

Previous work in the field has been developed along two routes: (1) predict the matrix elements (i.e., the hopping parameters) in the Hamiltonians of a set of similar material systems and create electronic band structures from the

Table 2 Results on the success rates (%) of band gap classification

	Method	Interaction	Atom	Atom+Interaction
Shallow	Decision tree	52.75	63.59	57.83
	Naive Bayes	57.03	57.45	61.78
	Multilayer perceptron	64.91	52.84	66.16
	SVM	67.89	55.40	66.23
	Random forest	66.16	64.80	67.84
GCN	MPNN w/o neighbor cell	–	–	69.12
	Cheb. Conv. w/o neighbor cell	–	–	70.39
	Cheb. Conv. w/ 1st neighbor cells	–	–	70.43

learned Hamiltonians [12, 25]; (2) learn from the electronic band structures and make predictions for the material Hamiltonians [42].

Unlike previous work, our work presents a general model that learns from the Hamiltonians for a large set of diverse inorganic materials with complex chemical formulas and predicts electronic structure-based materials properties (metal/non-metal classification) through a novel graph neural network framework.

Furthermore, the present approach is not limited to the metal/nonmetal classification. It opens new avenues for a broad range of scientifically and technologically critical applications such as topologically trivial/nontrivial quantum materials with diverse topological invariants, strong/weak light absorber materials for novel photovoltaics, and ideal/poor solid-state hosts of point defects for single-photon quantum emitters.

Acknowledgements NW and XQ gratefully acknowledge the support by the National Science Foundation (NSF) under award number OAC-1835690. Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing. Qimin Yan acknowledges the support by the U.S. Department of Energy, Office of Science, under award number DE-SC0020310.

Funding This research is supported in part by the NSF funding under award number OAC-1835690 and IIS-1814745, Department of Energy Office of Science funding under award number DE-SC0020310. The recent, present and anticipated employment of the authors are Temple University, Microsoft, Texas A&M University and Stony Brook University. There are no financial interests inferred.

Declarations

Conflict of interest The authors declare that there are no conflicts of interest.

References

1. Agrawal WLA, Choudhary A, Wolverton C (2016) A general-purpose machine learning framework for predicting properties of inorganic materials. npj Comput Mater 2(1):1–7
2. Bartok AP, Payne MC, Kondor R, Csanyi G (2010) Gaussian approximation potentials: the accuracy of quantum mechanics without the electrons. Phys Rev Lett 104:136403
3. Behler J (2011) Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations. Phys Chem Chem Phys 13:17930–17955
4. Bronstein MM, Bruna J, LeCun Y, Szlam A, Vandergheynst P (2017) Geometric deep learning: going beyond euclidean data. IEEE Sig Process Mag 34:18–42
5. Chandrasekaran A, Kamal D, Batra R, Kim C, Chen L, Ramprasad R (2019) Solving the electronic structure problem with machine learning. npj Comput Mater 5:22
6. Curtarolo S, Morgan D, Persson K, Rodgers J, Ceder G (2003) Predicting crystal structures with data mining of quantum calculations. Phys Rev Lett 91:135503
7. Defferrard M and Bresson X and Vandergheynst P (2016) Convolutional neural networks on graphs with fast localized spectral filtering. Adv Neural Inform Process Syste 3844 – 3852
8. Dwivedi VP and Joshi CK and Laurent T et al. (2020) Benchmarking graph neural networks 5:188 – 200
9. Fey M and Lenssen JE (2019) Fast graph representation learning with PyTorch Geometric
10. Ghiringhelli LM, Vybiral J, Levchenko SV, Draxl C, Scheffler M (2015) Big data of materials science: critical role of the descriptor. Phys Rev Lett 114:105503
11. Gilmer J and Schoenholz SS and Riley PF et al. (2017) Neural message passing for quantum chemistry. ICML, 1263 – 1272
12. Gu Q, Zhang L, and Feng J (2020) Neural network representation of electronic structure from ab initio molecular dynamics
13. Hamilton WL, Ying R, and Leskovec J (2017) Inductive representation learning on large graphs
14. Hansen K and Montavon G and Biegler et al. (2013) Assessment and validation of machine learning methods for predicting molecular atomization energies. J Chem Theory Comput 9:3404 – 3419
15. Hautier G, Fischer CC, Jain A, Mueller T, Ceder G (2010) Finding nature's missing ternary oxide compounds using machine learning and density functional theory. Chem Mater 22:3762–3767
16. Henaff M and Bruna J and LeCun Y (2015) Deep convolutional networks on graph-structured data
17. Anubhav J, Ping OS, Hautier Geoffroy C et al (2013) Commentary: the materials project: a materials genome approach to accelerating materials innovation. APL Mater 1:011002
18. Jain A, Ong SP, Hautier G, Chen W et al (2013) Commentary: the materials project: a materials genome approach to accelerating materials innovation. Apl Mater 1:011002
19. Kipf TN and Welling M (2016) Semi-supervised classification with graph convolutional networks. ICLR
20. Krizhevsky A and Sutskever I and Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inform Process Syst, 1097–1105
21. LeCun Y, Bottou L, Bengio Y, Haffner P (2008) Gradient-based learning applied to document recognition. Proc IEEE 86:2278–2324
22. Lee J, Seko A, Shitara K, Nakayama K, Tanaka I (2016) Prediction model of band gap for inorganic compounds by combination of density functional theory calculations and machine learning techniques. Phys Rev B 93(11):115104
23. Lee J, Seko A, Shitara K, Nakayama K, Tanaka I (2016) Prediction model of band gap for inorganic compounds by combination of density functional theory calculations and machine learning techniques. Phys Rev B 93:115104
24. LeSar R (2009) Materials informatics: an emerging technology for materials development. Stat Anal Data Min 1:372–374
25. Li H, Wang Z, Zou N, Ye M, Duan W, and Xu Y (2021) Deep neural network representation of density functional theory hamiltonian
26. Long CJ, Hattrick-Simpers J, Murakami M, Srivastava RC, Takeuchi I, Karen VL, Li X (2007) Rapid structural mapping of ternary metallic alloy systems using the combinatorial approach and cluster analysis. Rev Sci Instrum 78:072217
27. Marzari N, Vanderbilt D (1997) Maximally localized generalized Wannier functions for composite energy bands. Phys Rev B 56:12847–12865
28. Morawietz T, Behler J (2013) A density-functional theory-based neural network potential for water clusters including van der waals corrections. J Phys Chem A 117:7356–7366
29. Mostofi AA, Yates JR, Pizzi G et al (2014) An updated version of wannier90: a tool for obtaining maximally-localised Wannier functions. Comput Phys Commun 185:2309–2310

30. Olsthoorn B, Matthias Geilhufe R, Borysov SS, Balatsky AV (2019) Band gap prediction for large organic crystal structures with machine learning. Adv Quant Technol 2(7–8):1900023
31. Peter W and Hamrick JB and Bapst V and Sanchez-Gonzalez A et al. (2018) Relational inductive biases, deep learning, and graph networks
32. Pilania G, Wang C, Jiang X, Rajasekaran S, Ramprasad R (2013) Accelerating materials property predictions using machine learning. Sci Rep 3(1):1–6
33. Pilania G, Mannodi-Kanakkithodi A, Uberuaga BP, Ramprasad R, Gubernatis JE, Lookman T (2016) Machine learning bandgaps of double perovskites. Sci Rep 6:19375
34. Pilania G, Arun Mannodi-Kanakkithodi BP, Uberuaga Rampi Ramprasad, Gubernatis JE, Lookman T (2016) Machine learning bandgaps of double perovskites. Sci Rep 6(1):1–10
35. Pilania G, Gubernatis JE, Lookman T (2017) Multi-fidelity machine learning models for accurate bandgap predictions of solids. Comput Mater Sci 129:156–163
36. Qian X, Li J, Qi L et al (2008) Quasiatomic orbitals for ab initio tight-binding analysis. Phys Rev B 78:245112
37. Rajan K (2005) Materials informatics. Mater Today 8:38–45
38. Rajan AC, Mishra A, Satsangi S, Vaish R, Mizuseki H, Lee K-R, Singh AK (2018) Machine-learning-assisted accurate band gap predictions of functionalized MXene. Chem Mater 30(12):4031–4038
39. Shi Z, Tsymbalov E, Dao M, Suresh S, Shapeev A, Li J (2019) Deep elastic strain engineering of bandgap through machine learning. Proc Natll Acad Sci 116:4117
40. Snyder JC, Rupp M, Hansen K, Muller KR, Burke K (2012) Finding density functionals with machine learning. Phys Rev Lett 108:253002
41. Wu Z, Pan S, Chen F, Long G, Zhang C, and Yu Philip S (2020) A comprehensive survey on graph neural networks. IEEE Trans Neural Netw Learn sSyst
42. Ye WZS, Wang H, He J, Huang Q, Chang S (2021) Machine learning method for tight-binding Hamiltonian parameterization from ab-initio band structure. npj Comput Mater 7(1):1–10
43. Yuan D, Chuhan W, Zhang C, Liu Y, Cheng J, Lin J (2019) Bandgap prediction by deep learning in configurationally hybridized graphene and boron nitride. npj Comput Mater 5:26
44. Zhou J, Cui G, Zhang Z et al. (2018) Graph neural networks: a review of methods and applications
45. Zhuo Y, Tehrani AM, Brgoch J (2018) Predicting the band gaps of inorganic solids by machine learning. The J Phys Chem Lett 9(7):1668–1673