DFT Hamiltonian Neural Network Training with Semi-supervised Learning

Yucheol Cho*, Guenseok Choi*, Gyeongdo Ham
Mincheol Shin†, Dae-Shik Kim†
(*Equal contribution, †Equal correspondence)
Korea Advanced Institute of Science and Technology (KAIST)

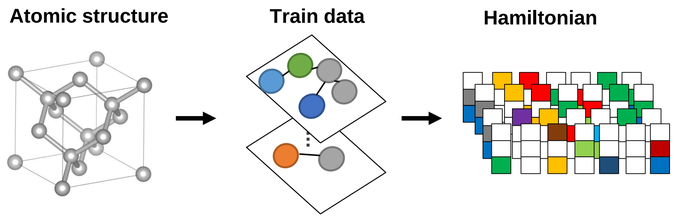
Introduction

- Previous neural network training methods required an extensive number of DFT simulations to obtain the ground truth (Hamiltonians).
- Conversely, when working with limited training data, deep learning models often exhibit increased errors in predicting Hamiltonians and band structures for testing data.
- This phenomenon carries the potential risk of yielding inaccurate physical interpretations, including the emergence of unphysical branches within band structures.
- To address this challenge, we introduce a novel deep learning-based method for calculating DFT Hamiltonians, specifically designed to generate accurate results with limited training data.

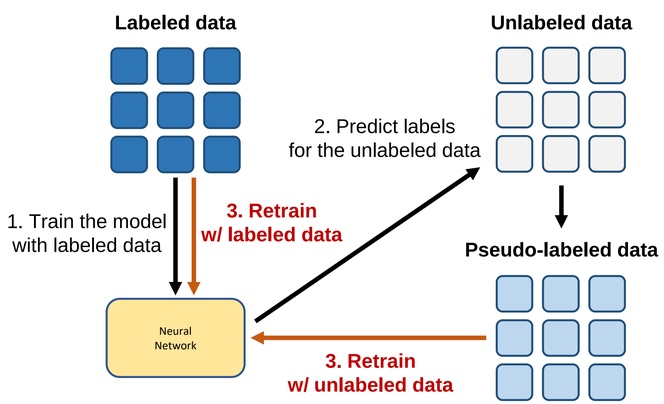
Preliminaries

Neural Network Hamiltonian

- DFT's Hamiltonian Message-passing Neural Networks streamline atomic structure analysis, offering an efficient alternative to traditional methods.



Semi-supervised Learning (Pseudo-label method)



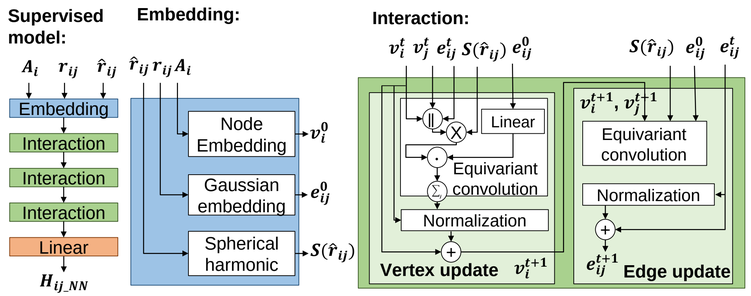
Method

Data Preparation

- Data preparation: atomic structure creation (Vienna ab-initio simulation package), DFT Hamiltonian calculation (OpenMX software), and Hamiltonian matrix transformation (Wigner D-matrix).
- Generated data for MoS2, Bi2Te3, HfO2, InGaAs.

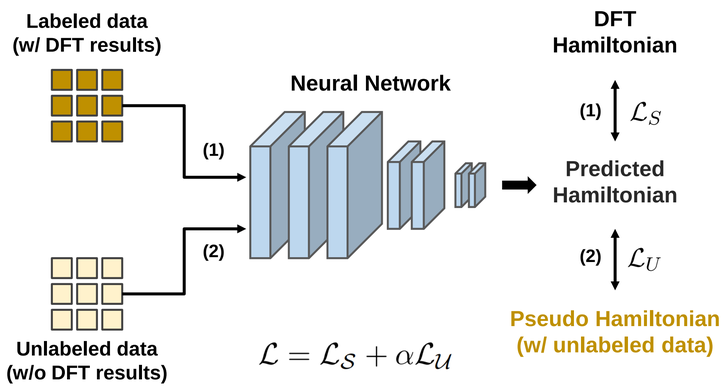
Message-passing Neural Network

- Our framework has been applied to the state-of-the-art model, DeepH-E3, which utilizes a message passing neural network.
- v_i is vertex i representing an atom, e_{ij} is the edge between vertices i and j indicating the bond, A_i is atomic number, r_{ij} is the distance, $\hat{r}_{ij}$ is the angle, H_{ij} is the Hamiltonian element, and t is the number of update.



Semi-supervised Learning

- We utilize mean-squared error loss for both supervised and unsupervised loss.
- Total loss includes the both losses with a hyperparameter **α**, which controls effects of unlabeled data for training.

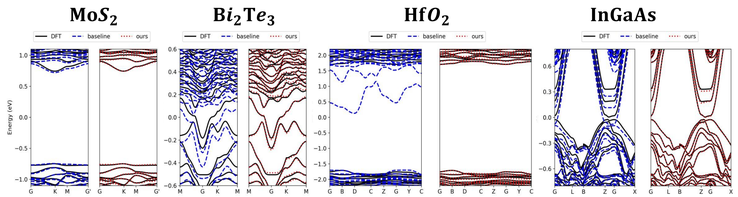


$$\mathcal{L} = \mathcal{L}_S + \alpha \mathcal{L}_U$$

Result

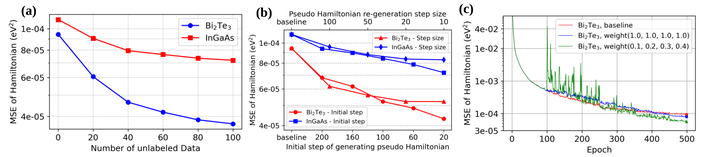
Band Structure

- Limited labeled data leads to inaccurate band structures; our approach with unlabeled data in neural network training corrects this.



Ablation Study

- More unlabeled data reduces Hamiltonian error (Figure a).
- Reducing initial step I and increasing training weight α gradually improve accuracy (Figure b and c).



Conclusion

- When dealing with a limited amount of training data, there exists a risk of obtaining distorted results in the subsequent physical analysis.
- To address this challenge, we introduce a framework that mitigates the limitations arising from insufficient training data.
- We achieve this by incorporating semi-supervised learning techniques into neural network training.

Broader Impact

- There are scenarios where access to extensive simulations or conducting experiments can be limited due to various constraints, including resource limitations or high costs.
- In such situations, it becomes imperative to achieve reliable and meaningful results using a limited amount of training data.
- Our framework offers a versatile solution that can be applied effectively to a wide range of examples in these circumstances.

https://nips.cc/virtual/2023/76093