CHGCNN: Crystal Hypergraph Neural Networks

A Universal Framework for Material System Representations

Alex Heilman, Weiyi Gong, Qimin Yan

Department of Physics, Northeastern University

March 5, 2024

Overview

- Crystal graphs

Overview

- Crystal graphs, MPNNs

Overview

- Crystal graphs, MPNNs, limitations

Overview

- Crystal graphs, MPNNs, limitations
- Crystal hypergraphs

Overview

- Crystal graphs, MPNNs, limitations
- Crystal hypergraphs, hyperedge types

Overview

- Crystal graphs, MPNNs, limitations
- Crystal hypergraphs, hyperedge types, MPNN extensions

Overview

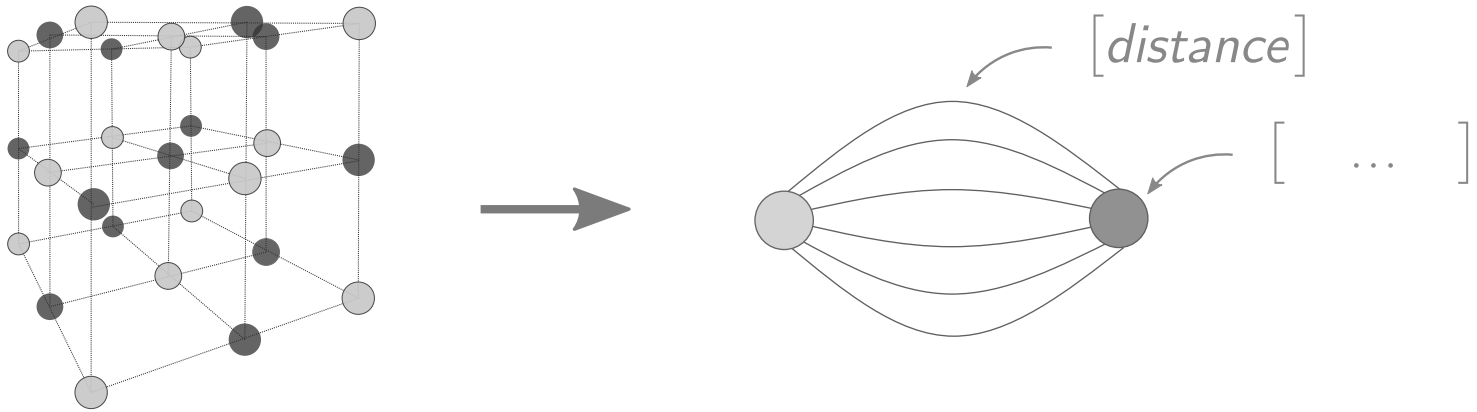
- Crystal graphs, MPNNs, limitations
- Crystal hypergraphs, hyperedge types, MPNN extensions
- See them in action!

Overview

- Crystal graphs, MPNNs, limitations
- Crystal hypergraphs, hyperedge types, MPNN extensions
- See them in action! Comparative testing

Crystal Graph Construction

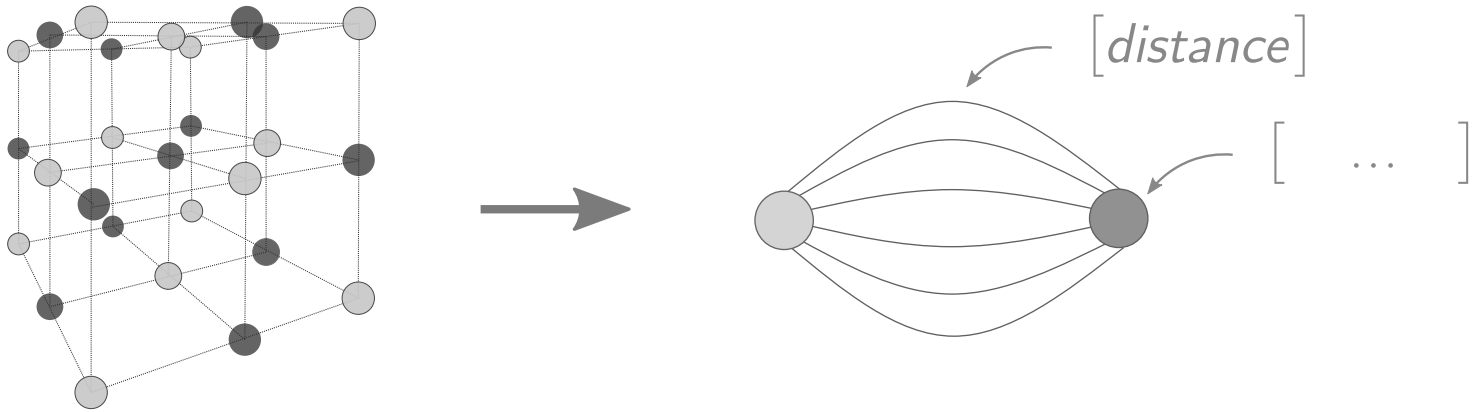
Usually represent crystalline systems as graphs (nodes and edges)[1]:



[1] Xie and Grossman, "Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties".

Crystal Graph Construction

Usually represent crystalline systems as graphs (nodes and edges)[1]:



- Atoms are nodes
- Edges denote 'neighbors'; based on distance
- Feature vectors

[1] Xie and Grossman, "Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties".

Message Passing on Graphs

MPNN framework[2]:

[2] Gilmer et al., "Neural message passing for quantum chemistry".

Message Passing on Graphs

MPNN framework[2]:

- Message m for each node i:

$$m_i^{t+1} = \sum_{n_j \in \mathcal{N}(i)} M_t(n_i^t, e_{ij}, n_j^t)$$

- Update function U for layer t:

$$n_i^{t+1} = U_t(n_i^t, m_i^{t+1})$$

[2] Gilmer et al., "Neural message passing for quantum chemistry".

Message Passing on Graphs

MPNN framework[2]:

- Message m for each node i:

$$m_i^{t+1} = \sum_{n_j \in \mathcal{N}(i)} M_t(n_i^t, e_{ij}, n_j^t)$$

- Update function U for layer t:

$$n_i^{t+1} = U_t(n_i^t, m_i^{t+1})$$

We'll need to generalize this later...

[2] Gilmer et al., "Neural message passing for quantum chemistry".

Graph Limitations

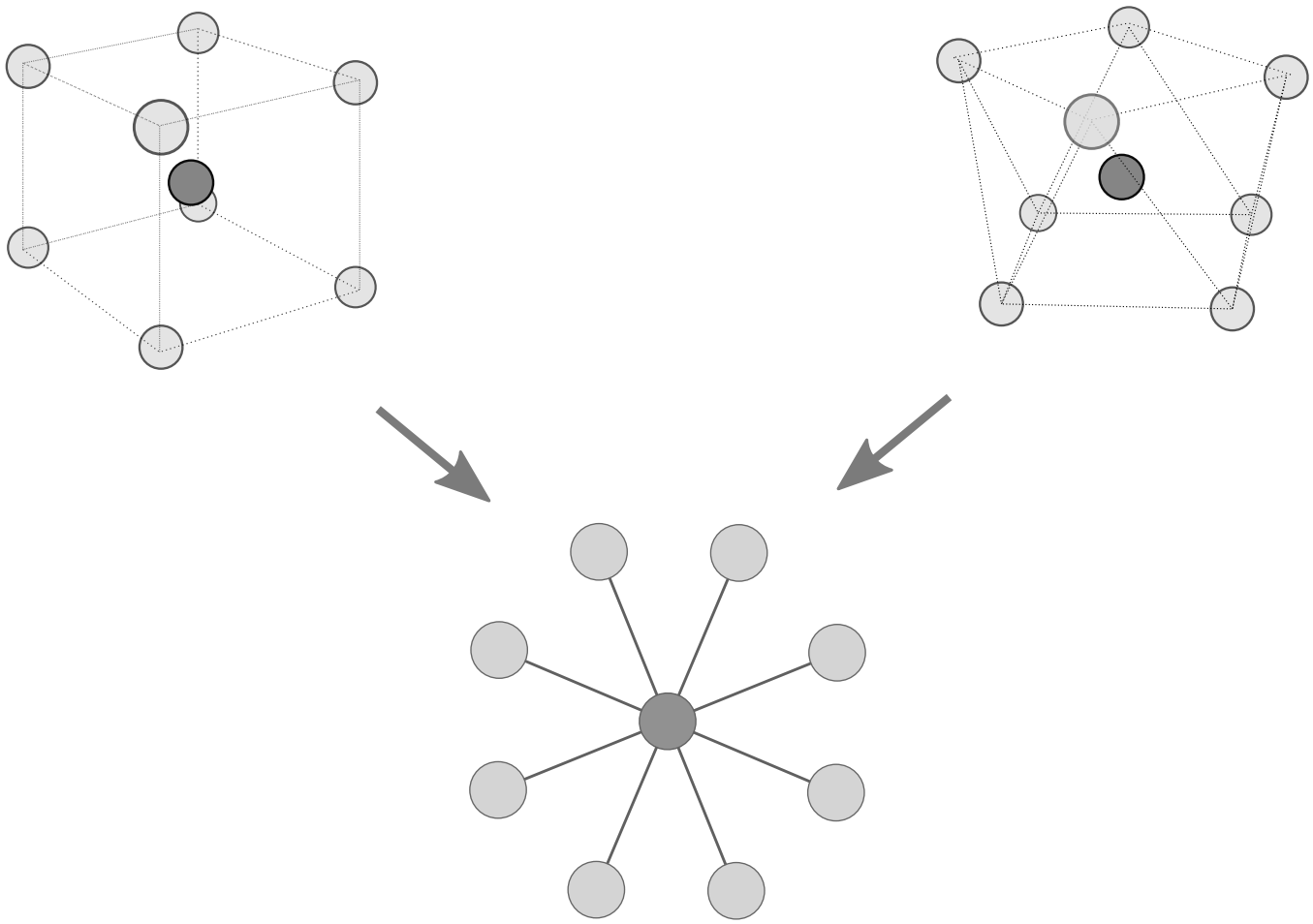
Problem:

Graph Limitations

Problem: Graph encodes ONLY interatomic distances!

Graph Limitations

Problem: Graph encodes ONLY interatomic distances!



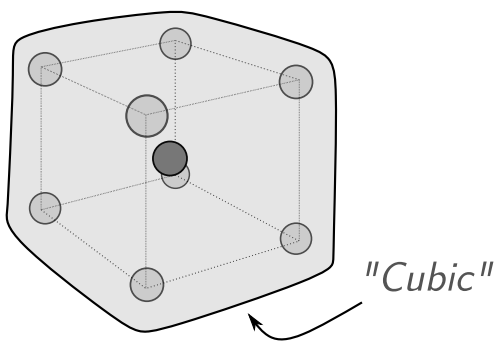
These two crystalline structures have same representations!

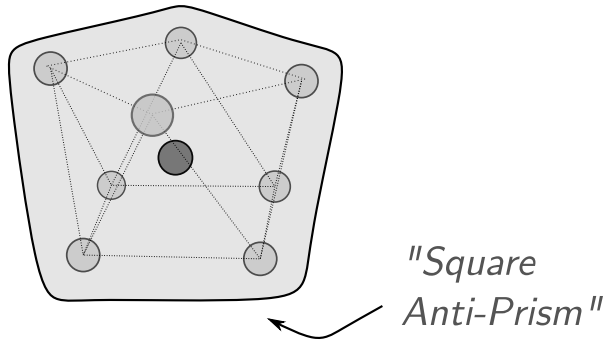
Solution: Hypergraphs!

Hypergraphs allow us to have larger hyperedges.

Solution: Hypergraphs!

Hypergraphs allow us to have larger hyperedges.

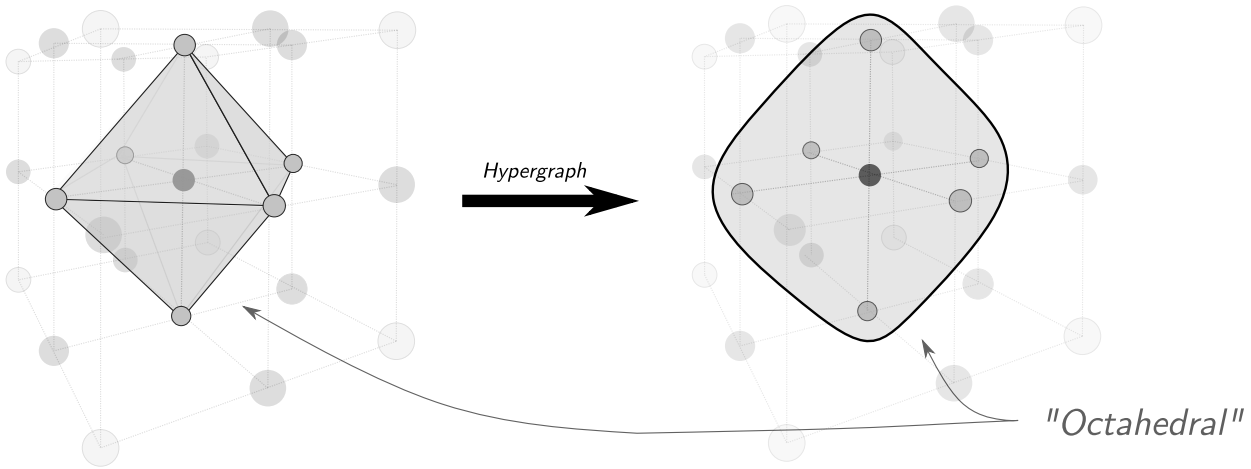




So we may define hyperedges for this local geometry.

Local Environments as Hyperedges

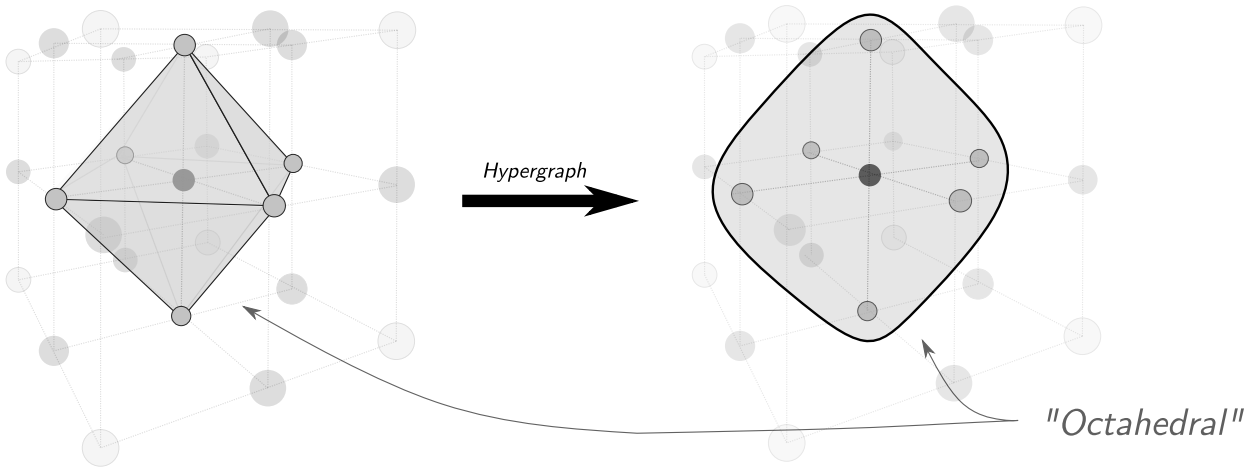
Construct a hyperedge for each atom's entire local environment:



[3] Zimmermann and Jain, "Local structure order parameters and site fingerprints for quantification of coordination environment and crystal structure similarity".

Local Environments as Hyperedges

Construct a hyperedge for each atom's entire local environment:

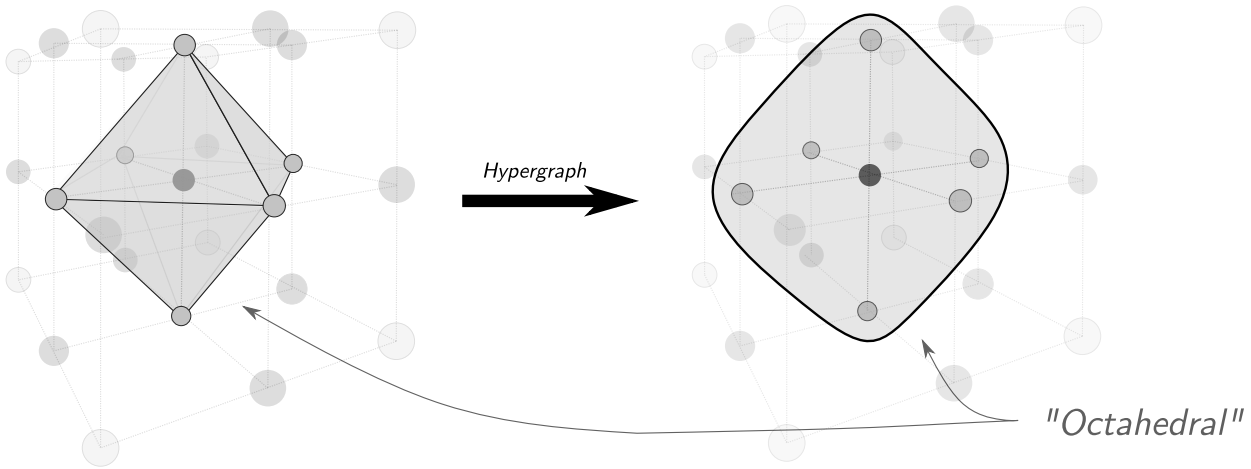


These contain the entire first shell of neighbors.

[3] Zimmermann and Jain, "Local structure order parameters and site fingerprints for quantification of coordination environment and crystal structure similarity".

Local Environments as Hyperedges

Construct a hyperedge for each atom's entire local environment:



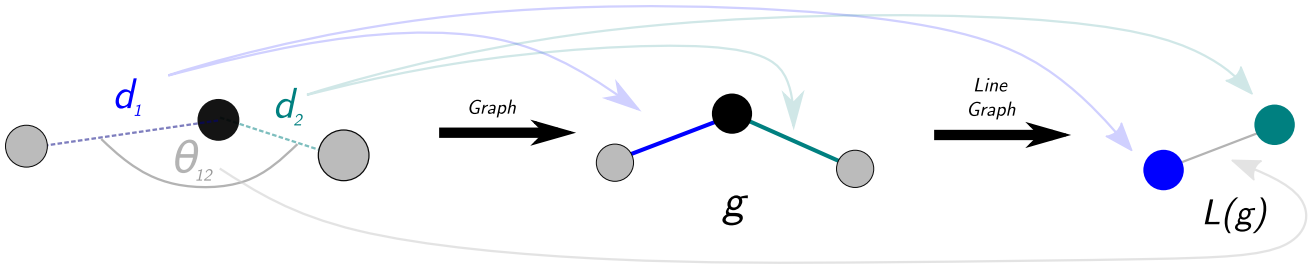
These contain the entire first shell of neighbors.

• Geometry is encoded quantitatively with continuous symmetry measures or structure order parameters[3] as features

[3] Zimmermann and Jain, "Local structure order parameters and site fingerprints for quantification of coordination environment and crystal structure similarity".

Another Example: Triplets as Hyperedges

Many modern models utilize bond angle information[4].

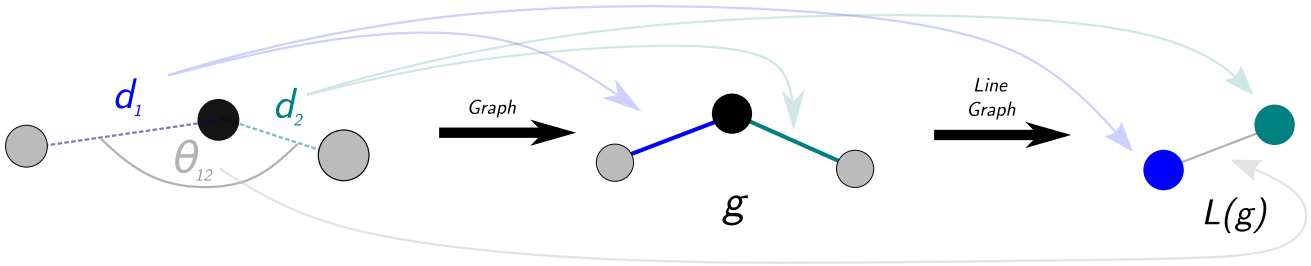


This is edge feature in a derived, auxiliary, 'line-graph'.

[4] Choudhary and DeCost, "Atomistic Line Graph Neural Network for improved materials property predictions"; Chen and Ong, "A universal graph deep learning interatomic potential for the periodic table".

Another Example: Triplets as Hyperedges

Many modern models utilize bond angle information[4].



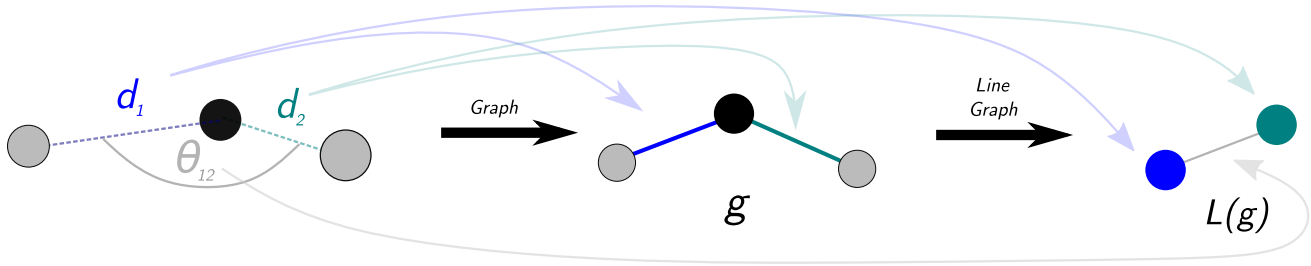
This is edge feature in a derived, auxiliary, 'line-graph'.

For hypergraphs, this is simpler:

[4] Choudhary and DeCost, "Atomistic Line Graph Neural Network for improved materials property predictions"; Chen and Ong, "A universal graph deep learning interatomic potential for the periodic table".

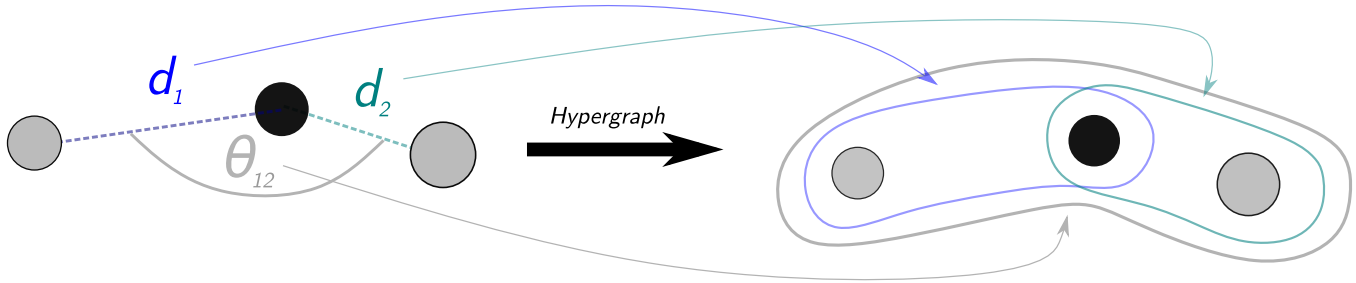
Another Example: Triplets as Hyperedges

Many modern models utilize bond angle information[4].



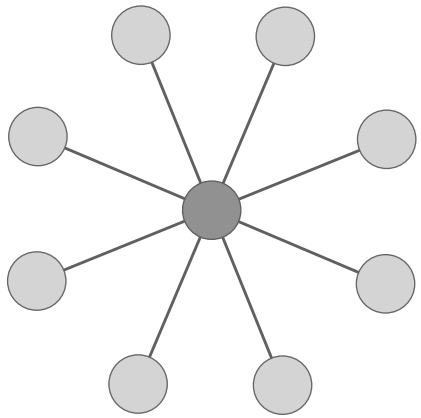
This is edge feature in a derived, auxiliary, 'line-graph'.

For hypergraphs, this is simpler:



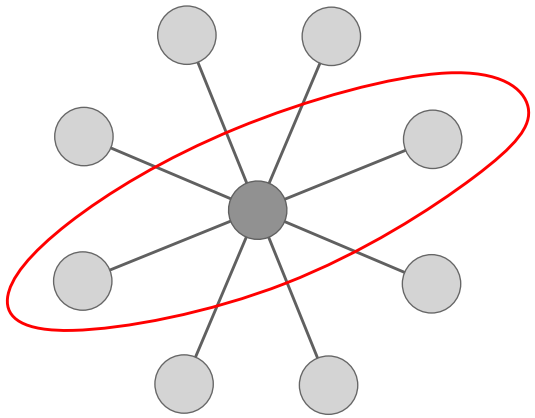
[4] Choudhary and DeCost, "Atomistic Line Graph Neural Network for improved materials property predictions"; Chen and Ong, "A universal graph deep learning interatomic potential for the periodic table".

Crystal Hypergraphs



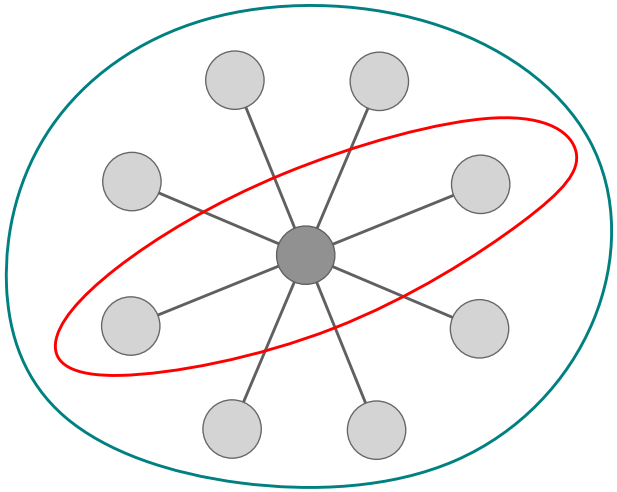
- In crystal hypergraph, all different order structures are on equal footing: bonds

Crystal Hypergraphs



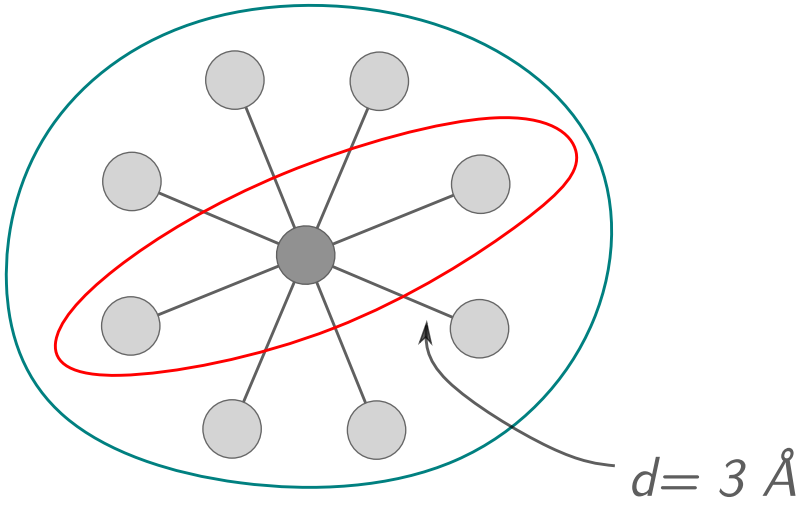
- In crystal hypergraph, all different order structures are on equal footing: bonds, triplets

Crystal Hypergraphs



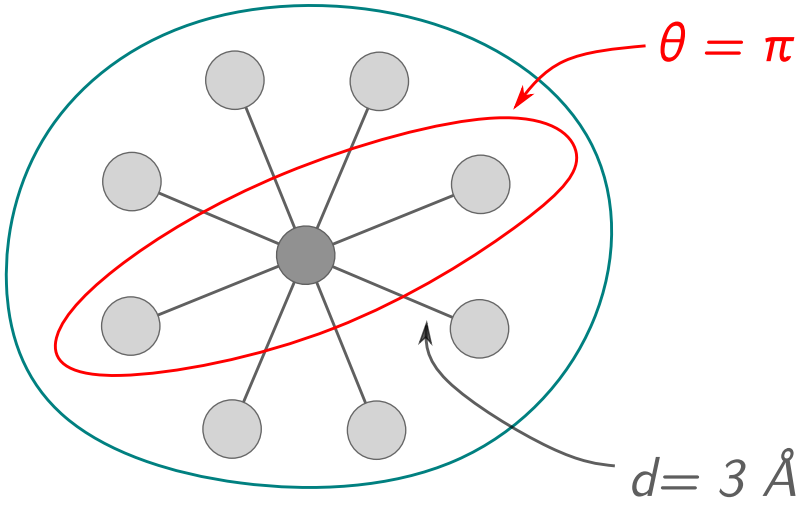
- In crystal hypergraph, all different order structures are on equal footing: bonds, triplets, motifs

Crystal Hypergraphs



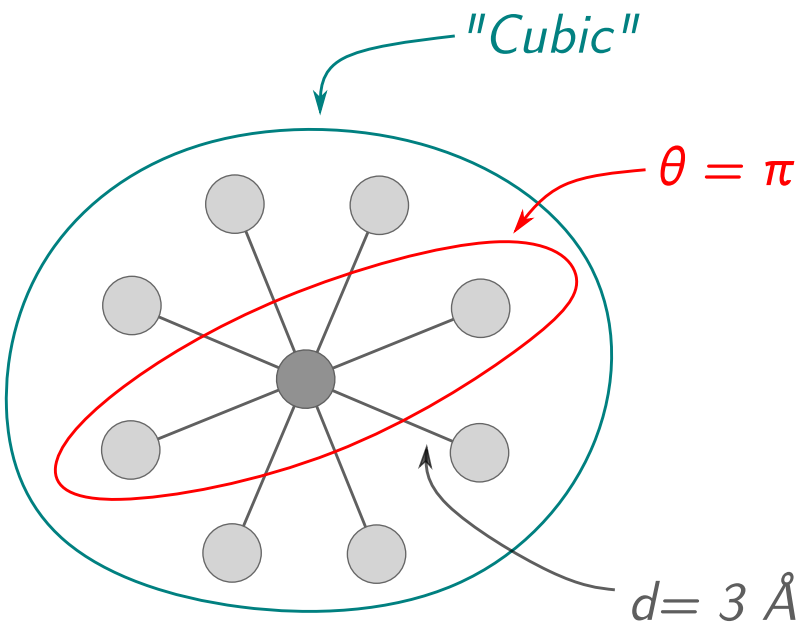
- In crystal hypergraph, all different order structures are on equal footing: bonds, triplets, motifs
- Each has hyperedge type with invariant features: distance

Crystal Hypergraphs



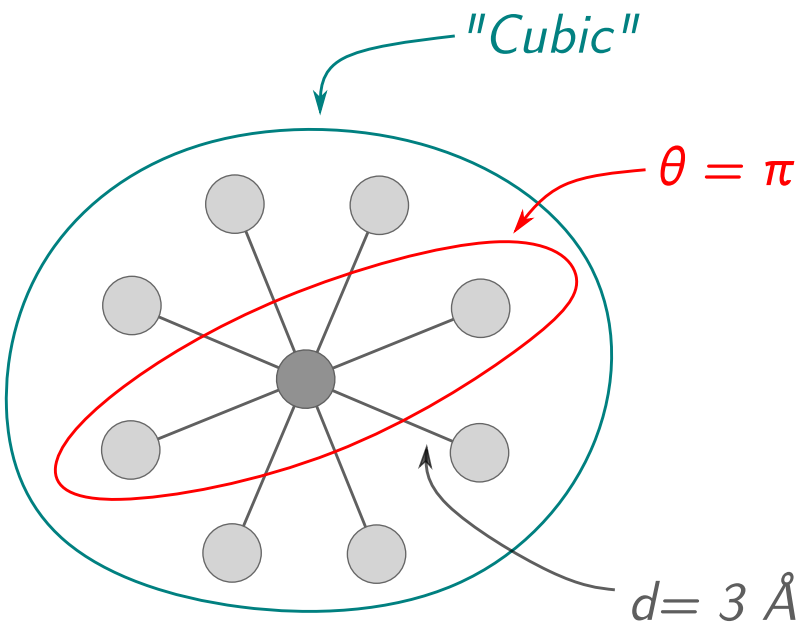
- In crystal hypergraph, all different order structures are on equal footing: bonds, triplets, motifs
- Each has hyperedge type with invariant features: distance, angle

Crystal Hypergraphs



- In crystal hypergraph, all different order structures are on equal footing: bonds, triplets, motifs
- Each has hyperedge type with invariant features: distance, angle, shape

Crystal Hypergraphs



- In crystal hypergraph, all different order structures are on equal footing: bonds, triplets, motifs
- Each has hyperedge type with invariant features: distance, angle, shape

But..... how do we update these representations?

Extending Message Passing to Hypergraphs

In the case of hypergraphs, neighborhood of nodes relevant to each message is now a set:

Extending Message Passing to Hypergraphs

In the case of hypergraphs, neighborhood of nodes relevant to each message is now a set:

$$m_i^{t+1} = \sum_{h_j \ni n_i} M_t(n_i^t, \underbrace{h_j^t, \{n_w^t | n_w \in h_j\}}_{e_{ij}, n_j})$$

$$n_i^{t+1} = U_t(n_i^t, m_i^{t+1})$$

Extending Message Passing to Hypergraphs

In the case of hypergraphs, neighborhood of nodes relevant to each message is now a set:

$$m_i^{t+1} = \sum_{h_j \ni n_i} M_t(n_i^t, \underbrace{h_j^t, \{n_w^t | n_w \in h_j\}}_{e_{ij}, n_j})$$

$$n_i^{t+1} = U_t(n_i^t, m_i^{t+1})$$

Three approaches are immediatley obvious.

Three Approaches to Hypergraph Convolution

1. Naive: One message for each hyperedge attribute

$$n_i' = n_i + f(n_i \oplus h_j)$$

Three Approaches to Hypergraph Convolution

1. Naive: One message for each hyperedge attribute

$$n_i' = n_i + f(n_i \oplus h_j)$$

2. Interorder: Message for each node in hyperedge

$$n_i' = n_i + \sum_{n_k \in h_j} f(n_i \oplus h_j \oplus n_k)$$

Three Approaches to Hypergraph Convolution

1. Naive: One message for each hyperedge attribute

$$n_i' = n_i + f(n_i \oplus h_j)$$

2. Interorder: Message for each node in hyperedge

$$n_i' = n_i + \sum_{n_k \in h_j} f(n_i \oplus h_j \oplus n_k)$$

3. Hyperedge Aggregation: One message for each hyperedge

$$n_i' = n_i + f(n_i \oplus h_j \oplus \mathrm{AGG}\{n_k \in h_j\})$$

Three Approaches to Hypergraph Convolution

Suppose there is n hyperedges, average m nodes per hyperedge:

Three Approaches to Hypergraph Convolution

Suppose there is n hyperedges, average m nodes per hyperedge:

1. $n_i' = n_i + f(n_i \oplus h_j) \rightarrow \#$ of messages scales as $\mathcal{O}(nm)$

Three Approaches to Hypergraph Convolution

Suppose there is n hyperedges, average m nodes per hyperedge:

1. $n_i' = n_i + f(n_i \oplus h_j) \rightarrow$ # of messages scales as $\mathcal{O}(nm)$
2. $n_i' = n_i + \sum_{n_k \in h_j} f(n_i \oplus h_j \oplus n_k) \rightarrow$ scales as $\mathcal{O}(nm^2)$

Three Approaches to Hypergraph Convolution

Suppose there is n hyperedges, average m nodes per hyperedge:

1. $n_i' = n_i + f(n_i \oplus h_j) \rightarrow$ # of messages scales as $\mathcal{O}(nm)$
2. $n_i' = n_i + \sum_{n_k \in h_j} f(n_i \oplus h_j \oplus n_k) \rightarrow$ scales as $\mathcal{O}(nm^2)$
3. $n_i' = n_i + f(n_i \oplus h_j \oplus \text{AGG}\{n_k \in h_j\}) \rightarrow$ scales as $\mathcal{O}(nm)$!

Three Approaches to Hypergraph Convolution

Suppose there is n hyperedges, average m nodes per hyperedge:

1. $n_i' = n_i + f(n_i \oplus h_j) \rightarrow$ # of messages scales as $\mathcal{O}(nm)$

2. $n_i' = n_i + \sum_{n_k \in h_j} f(n_i \oplus h_j \oplus n_k) \rightarrow$ scales as $\mathcal{O}(nm^2)$

3. $n_i' = n_i + f(n_i \oplus h_j \oplus \text{AGG}\{n_k \in h_j\}) \rightarrow$ scales as $\mathcal{O}(nm)$!

Approach (**3.**) seems a good balance of message quality and scaling behaviour

Comparative Performance Testing

- We implement a basic convolutional structure amenable for (**3**)

Comparative Performance Testing

- We implement a basic convolutional structure amenable for (**3**)

- We focus on models with bond hyperedges and motif hyperedges.

Comparative Performance Testing

- We implement a basic convolutional structure amenable for (**3**)

- We focus on models with bond hyperedges and motif hyperedges.

- Acts as testbed for what features/order correlations are most important for task.

Hyperparameters and Model Architecture

For each convolutional structure, testing was done for a model with 3 convolutional layers, an initial learning rate of 0.01, hidden node features of dimension 64, and a hidden output layer of dimension 128 (Similar to CGCNN's architecture). The loss function utilized is MSE.

Datasets are split 80% for training and 20% for validation tests.

Materials Projet Datasets

Below are results for various regression and classifications tasks for 152,605 materials.

Formation Energy

Model	Best MAE (eV/Atom)
Bond-only	0.177
Motif-only	0.088
Bond & Motif	**0.074**

Band Gap

Model	Best MAE (eV)
Bond-only	0.315
Motif-only	0.387
Bond & Motif	**0.301**

Metal/Non-metal

Model	Best Accuracy
Bond-only	.837
Motif-only	.853
Bond & Motif	**.860**

MatBench Datasets

Below, we present results on validation sets for various MatBench target sets.

Model	Phonons (1,265) Best MAE (cm^{-1})	Refractive Indices (4,764) Best MAE
Bond-only	78.9	**.4891**
Motif-only	65.9	.5328
Bond & Motif	**59.0**	.5222

Model	Perovskites (18,829) Best MAE (eV/Atom)
Bond-only	0.329
Motif-only	**0.052**
Bond & Motif	0.058

Overview of Experimental Results

- Electronic tasks (band gap, dielectric targets) seem to benefit much less, or be negatively impacted, by this additional information.

Overview of Experimental Results

- Electronic tasks (band gap, dielectric targets) seem to benefit much less, or be negatively impacted, by this additional information. It seems pair-wise correlators are, in general, a better descriptor for such tasks.

Overview of Experimental Results

- Electronic tasks (band gap, dielectric targets) seem to benefit much less, or be negatively impacted, by this additional information. It seems pair-wise correlators are, in general, a better descriptor for such tasks.

- Formation energy tasks benefit greatly from included motif (i.e. higher order geometrical) descriptors

Overview of Experimental Results

- Electronic tasks (band gap, dielectric targets) seem to benefit much less, or be negatively impacted, by this additional information. It seems pair-wise correlators are, in general, a better descriptor for such tasks.

- Formation energy tasks benefit greatly from included motif (i.e. higher order geometrical) descriptors

- In the case of Perovskites (a harder set of formation energy targets), this geometrical information seems even more impactful

Overview of Experimental Results

- Electronic tasks (band gap, dielectric targets) seem to benefit much less, or be negatively impacted, by this additional information. It seems pair-wise correlators are, in general, a better descriptor for such tasks.

- Formation energy tasks benefit greatly from included motif (i.e. higher order geometrical) descriptors

- In the case of Perovskites (a harder set of formation energy targets), this geometrical information seems even more impactful. Though, it also may just help more for the smaller dataset available for training.

Overview of Experimental Results

• Electronic tasks (band gap, dielectric targets) seem to benefit much less, or be negatively impacted, by this additional information. It seems pair-wise correlators are, in general, a better descriptor for such tasks.

• Formation energy tasks benefit greatly from included motif (i.e. higher order geometrical) descriptors

• In the case of Perovskites (a harder set of formation energy targets), this geometrical information seems even more impactful. Though, it also may just help more for the smaller dataset available for training.

• Phonons (also heavily dependent on geometrical information) seem to also benefit greatly from motif-level features

Conclusion

- Graphs are limited in their expression of higher-order (above pair-wise) features of crystal structures

Conclusion

- Graphs are limited in their expression of higher-order (above pair-wise) features of crystal structures

- Crystal hypergraphs give us a natural way to encode different types of features in one representation

Looking Forward

- We may want to update nodes AND hyperedge features.

Looking Forward

- We may want to update nodes AND hyperedge features.
- Improve convolutional layers for hyperedges.

Looking Forward

- We may want to update nodes AND hyperedge features.
- Improve convolutional layers for hyperedges.
- Apply to novel tasks related to geometry.

Thanks for your time!

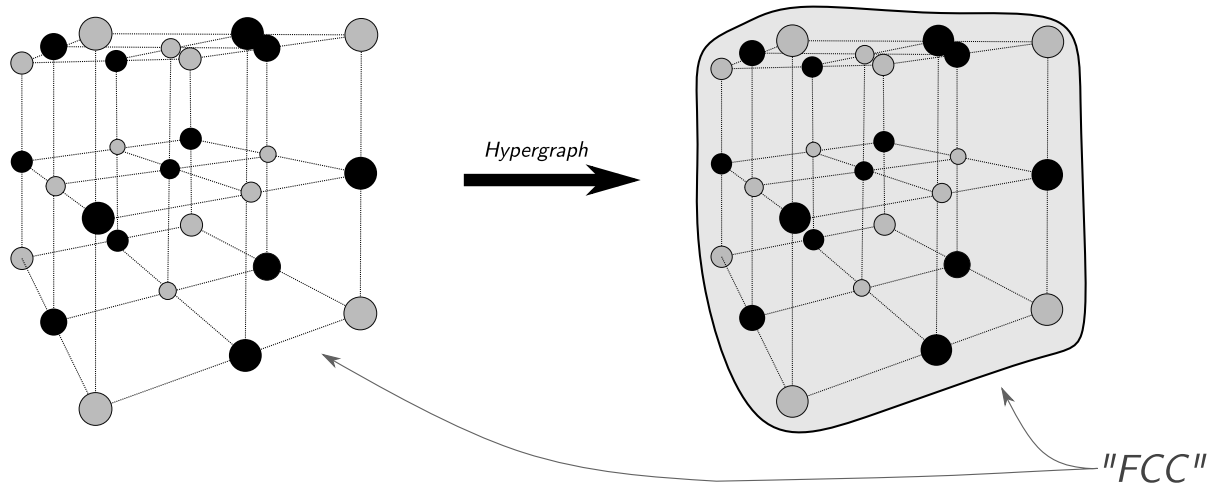
Hyperedge Index

Computationally, we define a *hyperedge index* of dimension 2 (as in previous hypergraph works):

[[node-index, ...],
 [hyperedge-index, ...]]

Cell Hyperedges

Another order of hyperedge we may consider is that describing the entire unit cell of some crystalline structure.

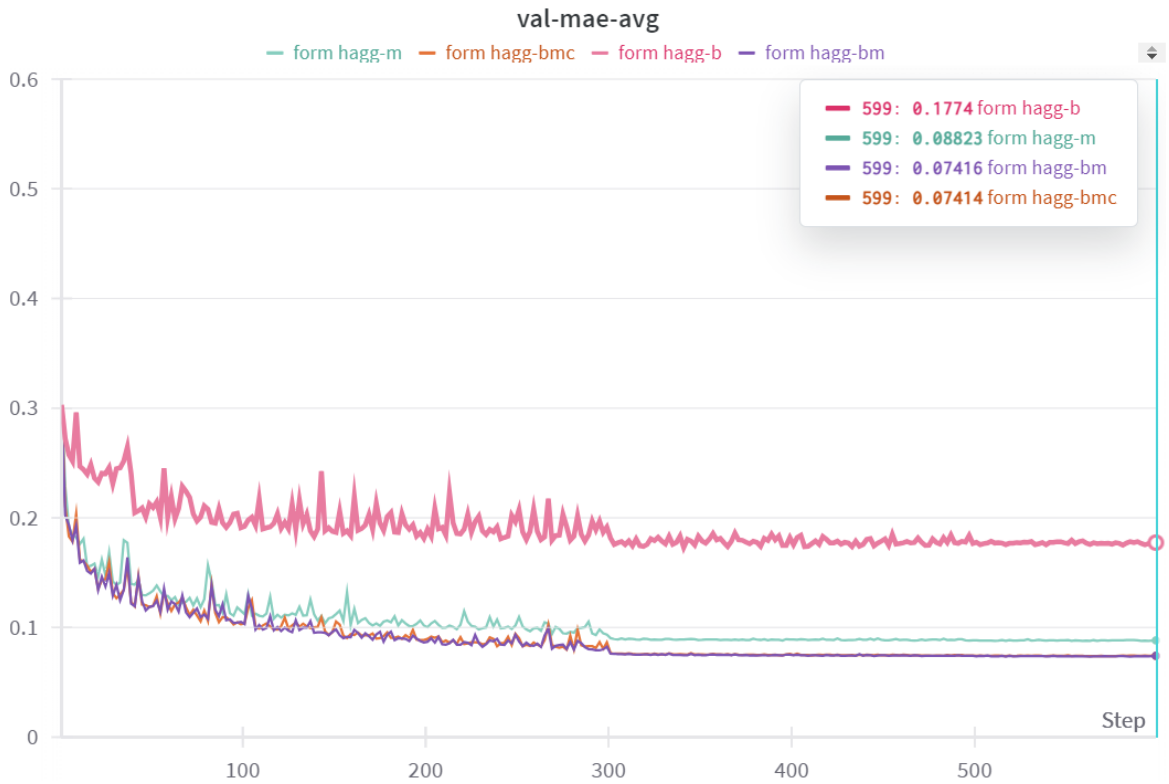


These 'unit' cell hyperedges allow for the explicit inclusion of global crystalline structure properties, i.e. point group information.

This unit cell feature also may be learned through convolution and potentially used as a dynamic state vector.

(Materials-Project) Formation Energy

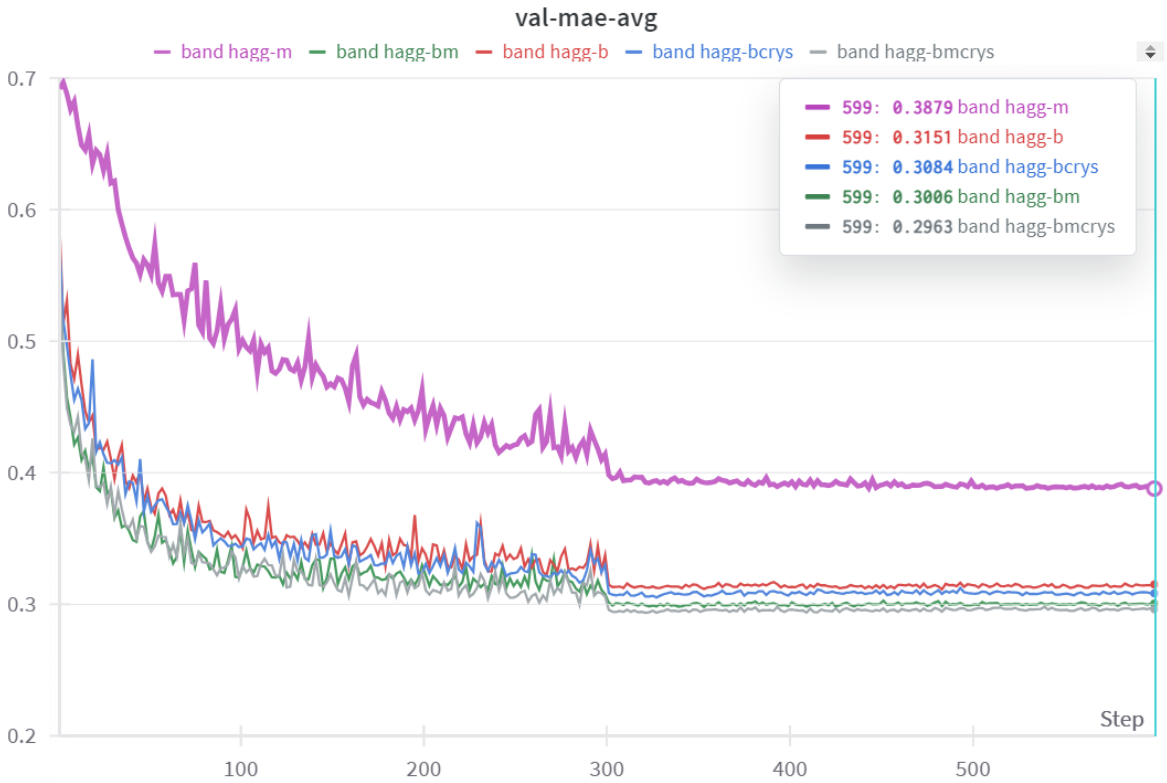
Below, validation MAE is shown through training for formation energy. The total dataset includes 152,605 materials.



Model	Best MAE (eV/Atom)
Bond-only	0.177
Motif-only	0.088
Bond & Motif	**0.074**

(Materials-Project) Band Gap

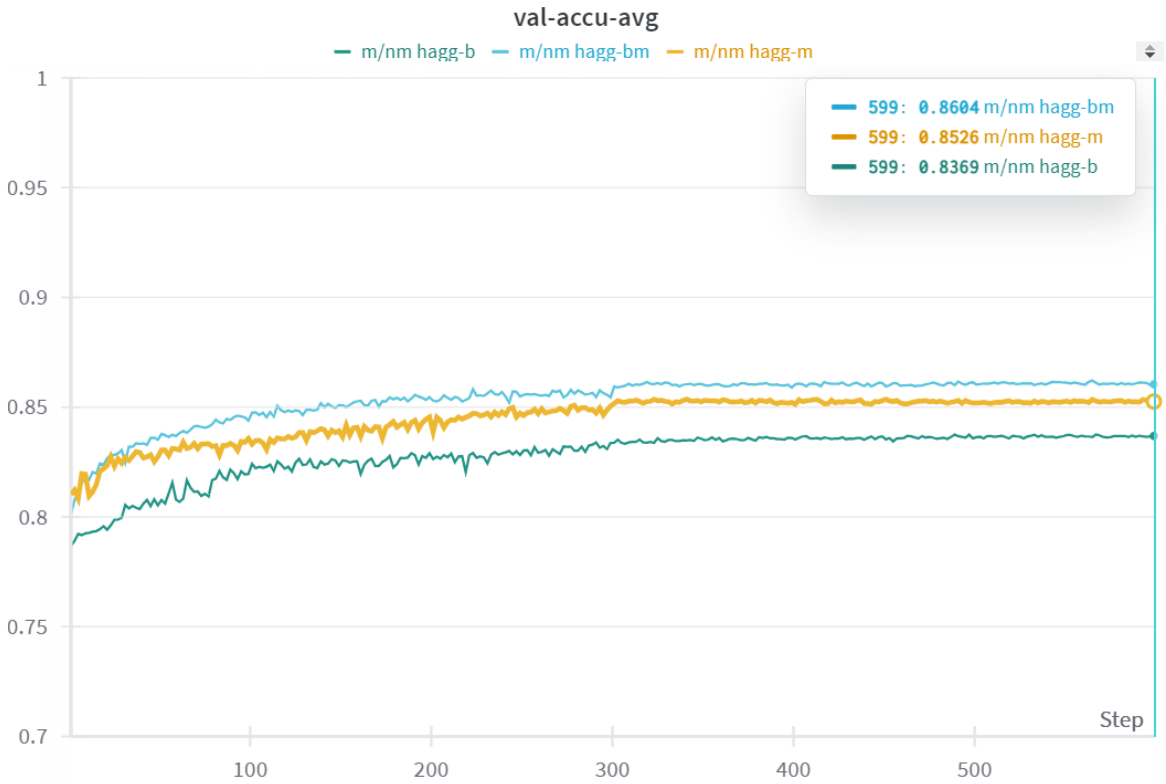
Below are results for band gap training, based on a dataset including 152,605 materials.



Model	Best MAE (eV)
Bond-only	0.315
Motif-only	0.387
Bond & Motif	**0.301**

(Materials-Project) Metal/Non-metal

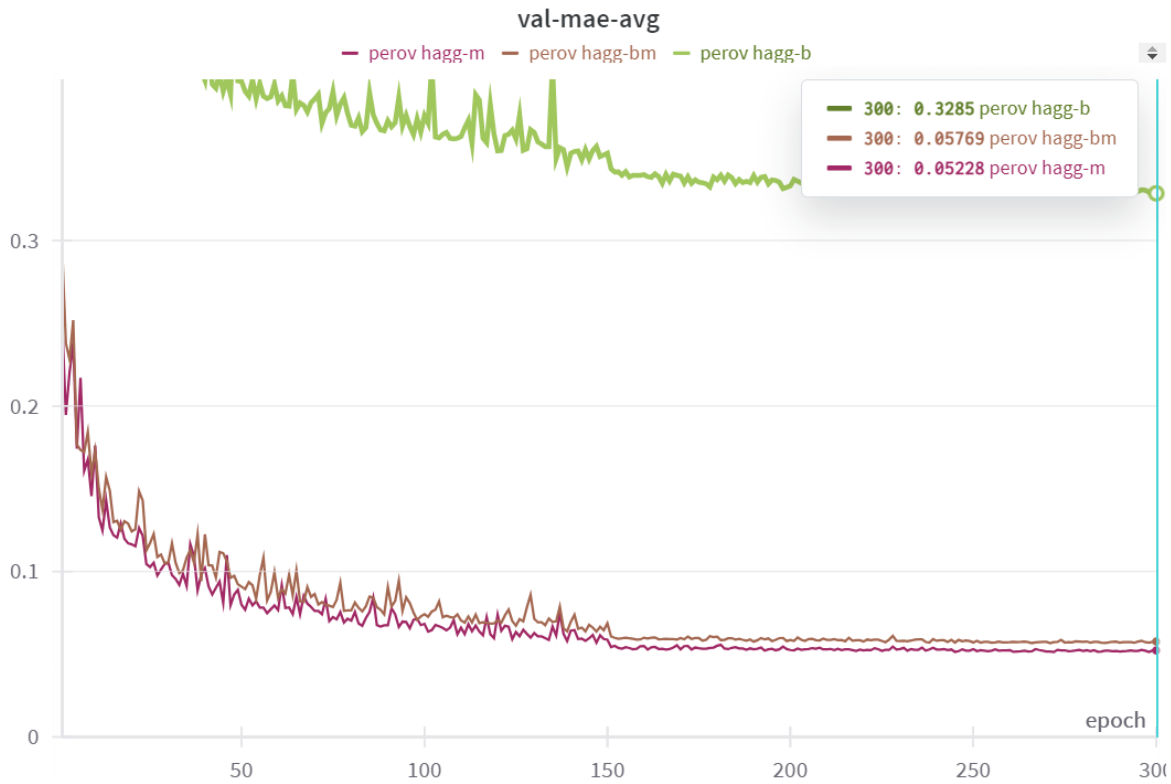
Below, we test metal/non-metal classification for 152,605 materials.



Model	Best Accuracy
Bond-only	.837
Motif-only	.853
Bond & Motif	**.860**

(MatBench) Perovskites

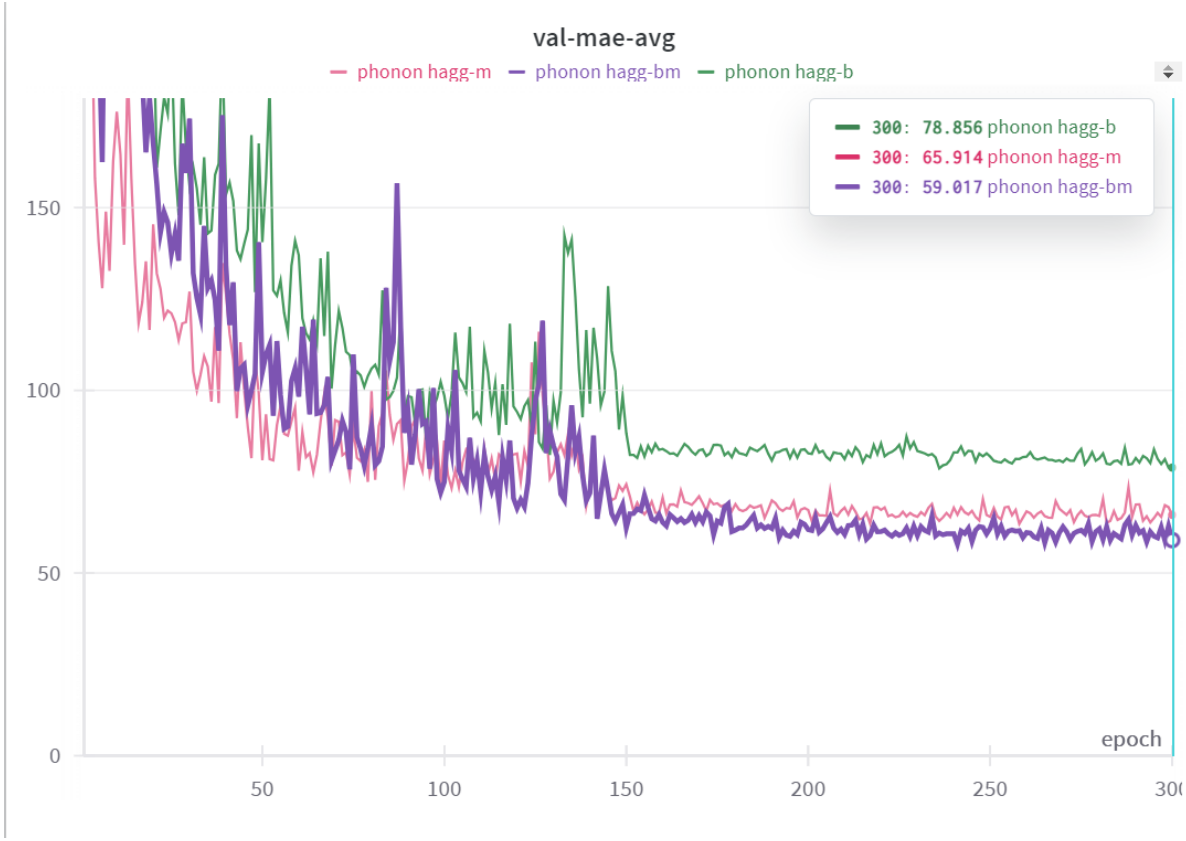
Below, we test on 18,928 calculated formation energies for Perovskites.



Model	Best MAE (eV/Atom)
Bond-only	0.329
Motif-only	**0.052**
Bond & Motif	0.058

(MatBench) Phonons

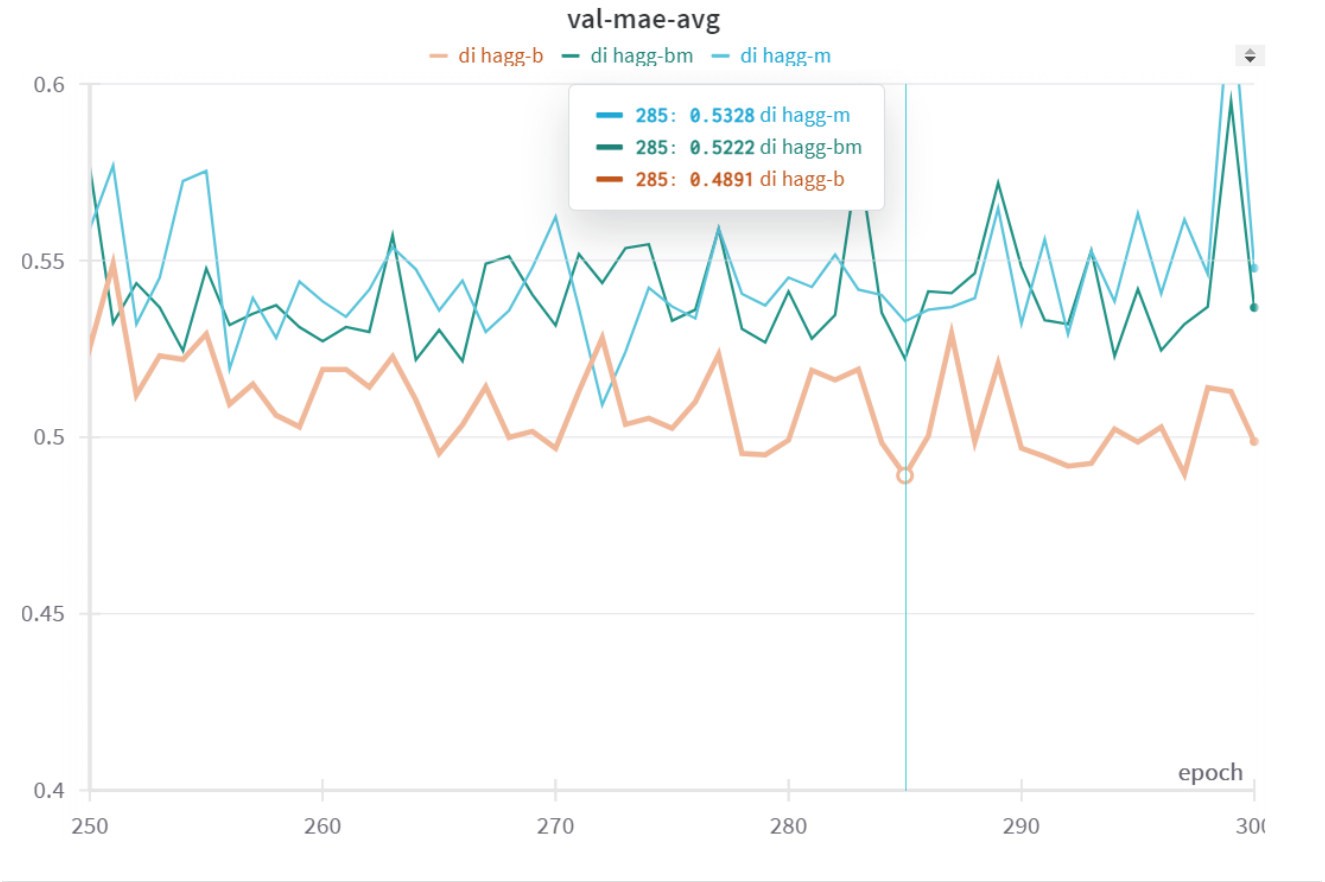
Below, we test on a dataset of 1,265 materials with the target being the highest calculated frequency optical phonon mode peak.



Model	Best MAE (cm^{-1})
Bond-only	78.9
Motif-only	65.9
Bond & Motif	**59.0**

(MatBench) Dielectrics

Below, we test on the calculated refractive index of 4,764 materials



Model	Best MAE
Bond-only	**.4891**
Motif-only	.5328
Bond & Motif	.5222