Crystal Hypergraph Networks

Project Proposal

Alexander Heilman

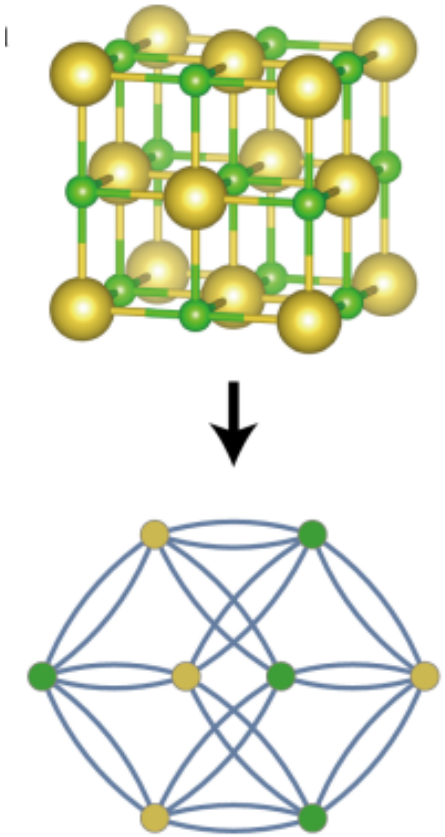
November 18, 2022

Crystal Graphs

- Crystalline structures may be represented as graphs (mathematical structure)

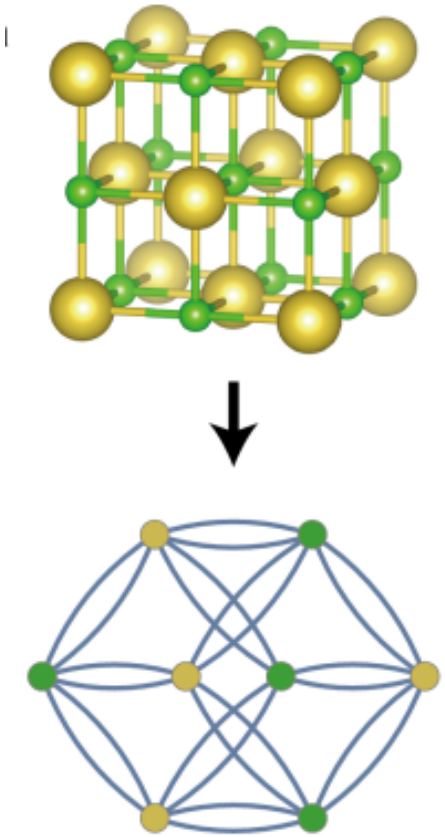
Crystal Graphs

- Crystalline structures may be represented as graphs (mathematical structure)
- Relevant features of the crystal structure's components may be associated with each node or edge



Crystal Graphs

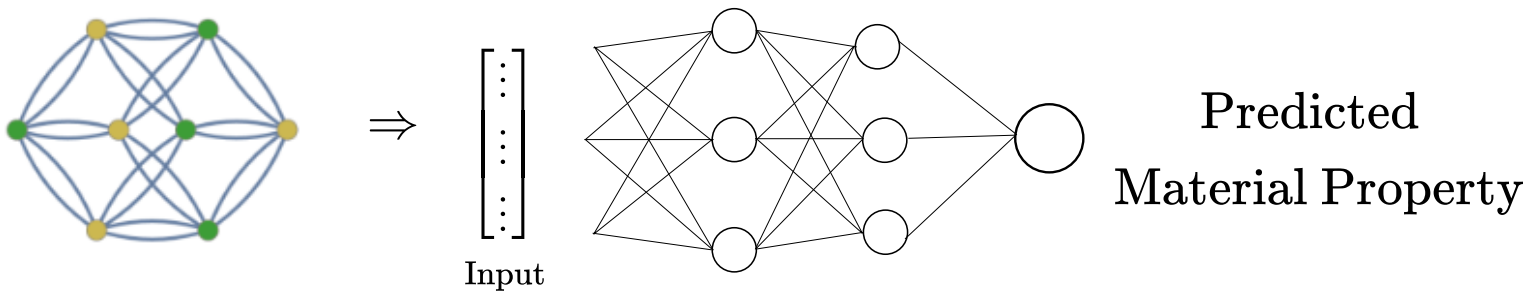
- Crystalline structures may be represented as graphs (mathematical structure)
- Relevant features of the crystal structure's components may be associated with each node or edge



So why do we care?

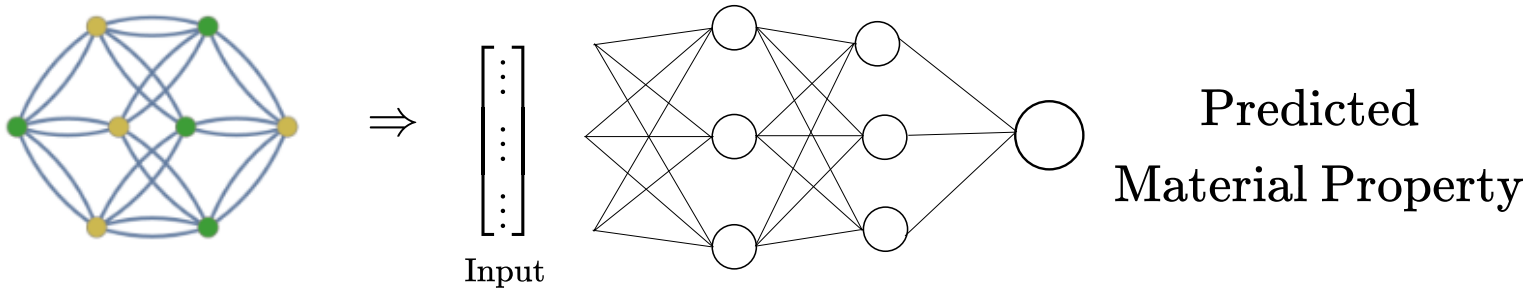
Crystal Graph Neural Networks

These graph structures may be used as input into graph neural networks (GNNs), which may then be trained to predict material properties.



Crystal Graph Neural Networks

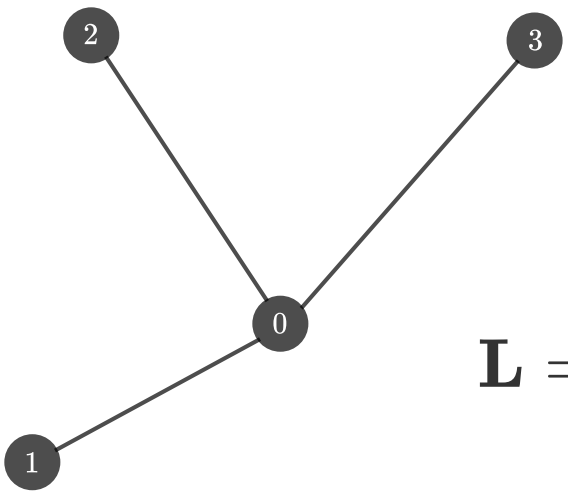
These graph structures may be used as input into graph neural networks (GNNs), which may then be trained to predict material properties.



- Formation Energies
- Bandgaps
- Bulk Moduli

Graph Convolution

Graph convolution may be defined as a product of matrices defined by the graph's structure and features.



$$\mathbf{L} = \mathbf{D} - \mathbf{A} =$$

Graph Laplacian

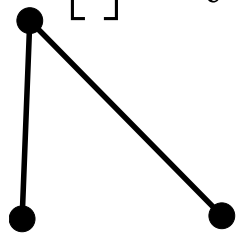
$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Graph Convolution (cont.)

A localized convolutional layer may be defined as a polynomial filter over the graph Laplacian [Welling and Kipf, 2016]:

$$\rho_\omega(L) = \sum_{i=0}^{d} \omega_i L^i$$

which is multiplied by the feature matrix of the graph, after which an activation function is applied (for each layer).

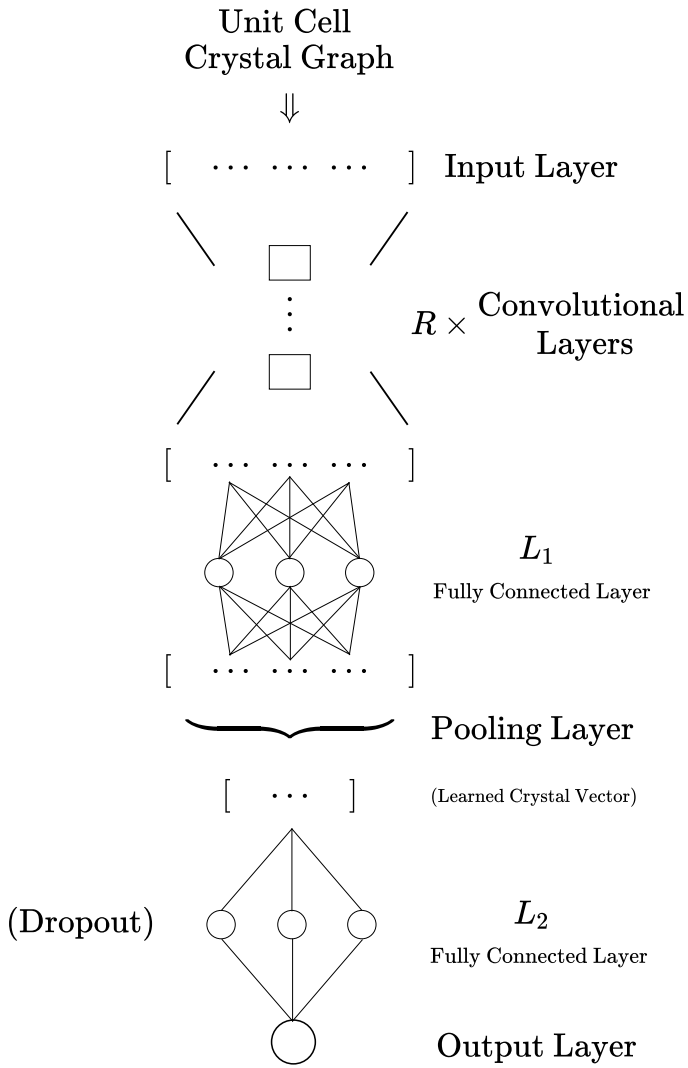


$$[\cdots|\cdots|\cdots]^{\mathbf{T}} = \mathbf{h}^{(k-1)}$$

$$\Downarrow$$

$$g\Big(\rho_\omega(L)[\cdots|\cdots|\cdots]^{\mathbf{T}}\Big) = \mathbf{h}^{(k)}$$

Example: Crystal Graph Convolutional Network



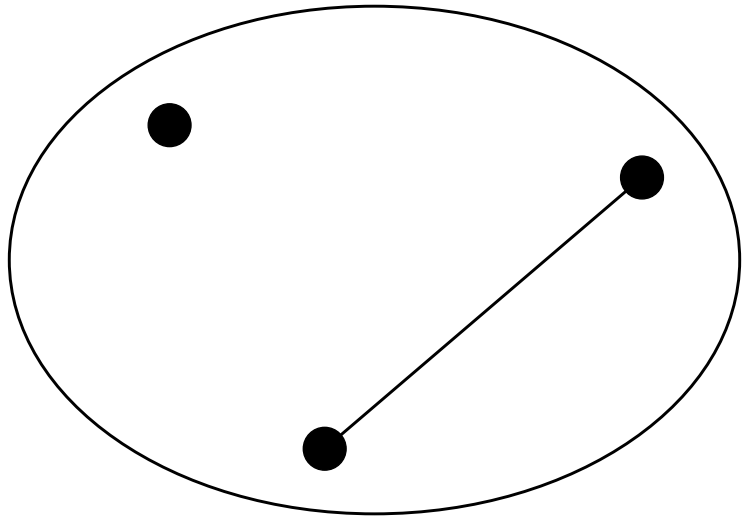
As a specific example, consider Xie and Grossman's Crystal Graph Convolutional Network [Xie and Grossman, 2018] (CGCNN) architecture.

This architecture was tested on crystalline structures for several targets with the following results:

Formation Energy: 0.038 eV/atom (MAE - 28046)

Band Gap: 0.388 eV (MAE - 16458 crystals)

Hypergraphs



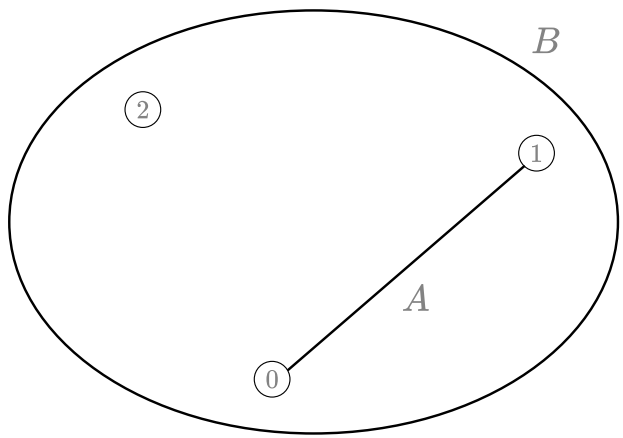
Hypergraphs relax the constraint of regular graphs restricting edges to pair-wise connections.

Hypergraph Convolution

Hypergraph convolution may be defined in a similar manner, as defined by Bai, Zhang & Torr [Bai et al., 2021a].

$$\mathbf{h}^{(k)} = g\left(\mathbf{H}\mathbf{W}\mathbf{H}^{\mathsf{T}}\mathbf{h}^{(\mathbf{k}-\mathbf{1})}\mathbf{P}\right)$$

where $\mathbf{H}$ is the hypergraph incidence matrix, $\mathbf{W}$ is the (optional, diagonal) hyperedge weight matrix, and $\mathbf{P}$ is the inter-layer weight matrix.



$$\mathbf{H} = \begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$

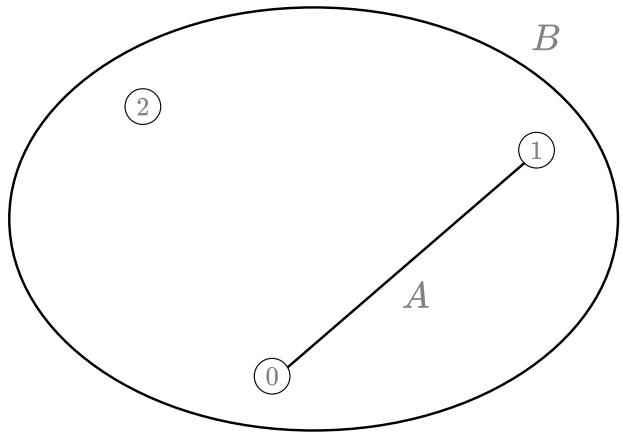
Incidence Matrix

Hypergraph Convolution

Hypergraph convolution may be defined in a similar manner, as defined by Bai, Zhang & Torr [Bai et al., 2021a].

$$\mathbf{h}^{(k)} = g\left(\mathbf{H}\mathbf{W}\mathbf{H}^{\mathbf{T}}\mathbf{h}^{(\mathbf{k}-\mathbf{1})}\mathbf{P}\right)$$

where $\mathbf{H}$ is the hypergraph incidence matrix, $\mathbf{W}$ is the (optional, diagonal) hyperedge weight matrix, and $\mathbf{P}$ is the inter-layer weight matrix.



$$\mathbf{H} = \begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$

Incidence Matrix

Note that I've neglected normalization techniques here

Crystal Hypergraph Networks

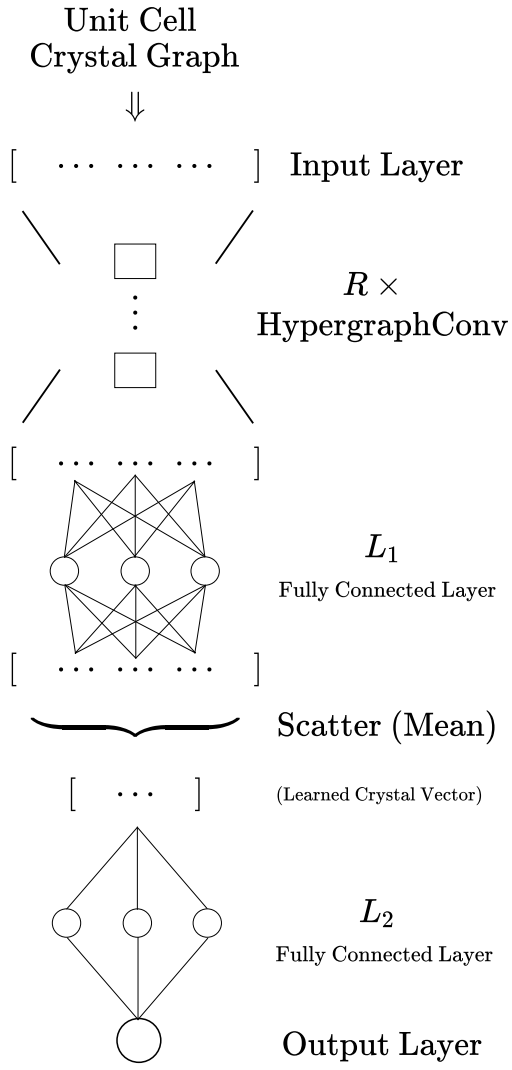
Since extra-pair-wise correlations are known to exist in crystal structures (a la mineral/oxide motifs, unit cells), we may leverage these higher order structures within this model to hopefully achieve better results.

Working Model

The current working model is (heavily) inspired by the CGCNN architecture. Convolutional layers are replaced with HypergraphConv layers in PyG [Bai et al., 2021b].

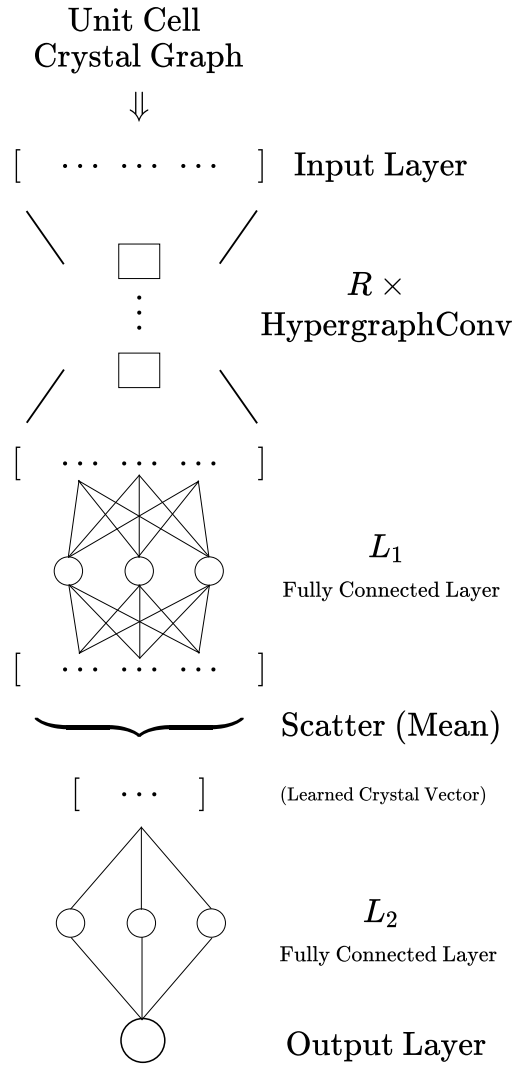
This architecture was tested on crystalline structures for several targets with the following results:

Formation Energy: 1 eV/atom (MAE) BAD!!



Working Model

The current working model is (heavily) inspired by the CGCNN architecture. Convolutional layers are replaced with HypergraphConv layers in PyG [Bai et al., 2021b].



This architecture was tested on crystalline structures for several targets with the following results:

Formation Energy: 1 eV/atom (MAE) BAD!!

CGCNN tested on same data had MAE around 0.1 eV/atom after similar epochs

Preliminary Results (BAD!)

- Edge/Hyperedge features lacking

Preliminary Results (BAD!)

- Edge/Hyperedge features lacking, how to generalize edge features to arbitrary order structures?

Preliminary Results (BAD!)

- Edge/Hyperedge features lacking, how to generalize edge features to arbitrary order structures?

Currently using CGCNN feature vectors for atoms and no edge properties. Should consider generalized edge features, and can look over motif fingerprints [Banjade et al., 2021] for inspiration.

Preliminary Results (BAD!)

- Edge/Hyperedge features lacking, how to generalize edge features to arbitrary order structures?

Currently using CGCNN feature vectors for atoms and no edge properties. Should consider generalized edge features, and can look over motif fingerprints [Banjade et al., 2021] for inspiration.

- Methods of crystal hypergraph construction

Preliminary Results (BAD!)

- Edge/Hyperedge features lacking, how to generalize edge features to arbitrary order structures?

Currently using CGCNN feature vectors for atoms and no edge properties. Should consider generalized edge features, and can look over motif fingerprints [Banjade et al., 2021] for inspiration.

- Methods of crystal hypergraph construction

Current model only makes hyperedges for groups of atoms within specified radius (for each node). Ideal construction should contain multiple orders of structures (edges, motifs & unit cell) as well as features of vectors in hyperedge construction (for example, only grouping cations with anions and vice-versa).

Looking Forward

- Extend to hypergraph attention
- Extend to dual hypergraphs to represent hyperedges instead of just nodes

Thanks!

Thanks for your time & attention!


Bai, S., Zhang, F., and Torr, P. H. (2021a).
Hypergraph convolution and hypergraph attention.
Pattern Recognition, 110:107637.


Bai, S., Zhang, F., and Torr, P. H. (2021b).
HypergraphConv Layer in Pytorch Geometric.
https://pytorch-geometric.readthedocs.io/en/latest/modules/nn.html#torch_geometric.nn.conv.HypergraphConv.
[Online; accessed 15-November-2022].


Banjade, H. R., Hauri, S., Zhang, S., Ricci, F., Gong, W., Hautier, G., Vucetic, S., and Yan, Q. (2021).
Structure motif–centric learning framework for inorganic crystalline systems.
Science advances, 7(17):eabf1754.


Welling, M. and Kipf, T. N. (2016).
Semi-supervised classification with graph convolutional networks.
In *J. International Conference on Learning Representations (ICLR 2017)*.


Xie, T. and Grossman, J. C. (2018).
Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties.
Physical review letters, 120(14):145301.