Equivariant Graph Neural Networks for Prediction of Tensor Material Properties of Crystals

Alexander J. Heilman*, Claire Schlesinger*, and Qimin Yan

Northeastern University

Abstract

Modern E(3)-equivariant networks have enabled accurate and efficient predictions of rotationally equivariant properties beyond scalar quantities. Tensorial quantities of crystalline systems, such as dielectric, piezoelectric, and elasticity tensors, are highly correlated with functionalities of complex materials but computationally expensive to produce *ab initio*. In this work, we demonstrate a comprehensive framework for building equivariant network architectures for tensorial property predictions by explicitly developing spherical harmonic decompositions of second- to fouth-order tensorial properties using their general symmetries. Three distinct E(3)-equivariant convolutional structures are developed and applied to predict the components of these decompositions, allowing us to predict the aforementioned tensorial quantities in an equivariant manner and compare the model performance. The models are trained using a large set of existing data in public material databases. Model performance for different tensorial components and the role of symmetry constraints are discussed. We further report the transferability of these predictive models between different tensorial target sets. The work presents a fundamental step toward the effective use of modern equivariant networks for the discovery and design of functional materials.

1 Introduction

Traditional machine learning methods applied to the material sciences have focused on the prediction of invariant, scalar properties of material systems[**mlreview1**, **mlreview1.5**, **mlreview2**, **mlreview2.25**, **mlreview2.5**, **mlreview3**]. The development of equivariant models enables the prediction of coordinate-system-dependent outputs in a well-defined manner [**tensorfieldnetworks**, **e3nn˙paper**]. Some applications of these equivariant networks neglect the direct prediction of directional (i.e. coordinate system-dependent) quantities and instead are used to predict invariant quantities to greater accuracy, from which directional quantities are then often derived [**nequip**, **magnet**]. Other applications have utilized the full capabilities of equivariant models to predict outputs associated

---

*These authors contributed equally to this work

with spherical harmonics [**deeph-e3**, **xpainn**], or use equivariant outputs to predict tensorial quantities of solid-state materials [**tensornet**, **painn**, **etgnn**, **yan2024equivcrystalprediction**, **MatTen**, **StrainNet**, **piezo-npj**]. However, these tensorial prediction works often give very brief descriptions of the harmonics decompositions used.

In rotationally equivariant models, features are often associated with irreducible representations of the rotation group [**e3nn˙paper**, **tensorfieldnetworks**], generally referred to as spherical harmonics. This association may be leveraged to predict tensorial quantities directly from the outputs of such models. This component-wise prediction of tensorial properties is achieved by decomposing tensors into harmonic subspaces via spherical harmonic decomposition, by which we may also associate arbitrary tensors with the irreducible representations of the rotation group [**harmonictensors**]. This allows one to essentially read off tensors component-wise from the output representations of these equivariant models by matching rotational and azimuthal indices.

Here, we give explicit harmonic decompositions of three material property tensors: the dielectric tensor; the piezoelectric tensor; and the elasticity tensor, following the approaches given in [**itin-elastic**, **itin-rank3**]. Results are then presented for for the prediction of these material property tensors directly from crystalline structures, as well as pretraining sequences composed of different scalar and tensorial targets investigating the transferability of such models between tasks. The structure of this work is as follows: section 2 gives a brief overview of the components of modern rotation-equivariant networks; section 3 details the harmonic decomposition of tensor spaces, drawing an analogy to feature spaces of equivariant models. The explicit decomposition of three common material property tensors is then given in section 4. Section 5 gives an overview of the basic model architecture that may be used to predict such tensors, as well as the specific implementations used here; and section 6 presents the prediction accuracy of such models in the prediction of the aforementioned material property tensors, as well as results detailing a set of pretraining chains between different target sets.

2 Rotationally Equivariant Neural Networks

Neural networks are a class of universal function approximators, composed layer-wise by functions $\mathcal{L}^1 \circ \mathcal{L}^2 \circ ... \circ \mathcal{L}^n$, where each layer-to-layer transition map $\mathcal{L}^i$ is a trainable function from some feature space associated with layer L to a feature space associated with layer $L + 1$.

A specific class of neural networks, referred to as Graph Neural Networks (GNN) [**cgcnn**, **alignn**, **amdnet**, **chemgnn**, **geocgcnn**, **icgcnn**] or, more generally, message passing networks (MPNN) [**mpnn**], act on feature vectors v_n^L, which are associated with nodes n in a representative graph (i.e. a collection of nodes and connections between them) for each layer L of the network. These models preserve the underlying connectivity of the graph representation, and act in a way that is invariant under node-index permutation.

MPNNs are generally invariant under translations as long as only coordinate-system invariant features (e.g. interatomic distance) are included. As such, these models are incapable of predicting any coordinate-system quantities, such as vector or tensor components. In material systems, we generally desire repre-

sentations that are equivariant under the group of transformations acting on real space. This corresponds to the special Euclidean group $SE(3)$, which consists of translations of the origin and the group of proper rotations $SO(3)$ [**e3nn˙paper**].

Equivariant functions are functions that preserve the action of a group on it's domain in it's codomain. For a function $f : X \to Y$, where X and Y are vector spaces with some set of group representations $\mathcal{D}^X(G)$ and $\mathcal{D}^Y(G)$ acting on these spaces respectively, f is equivariant with respect to G if it satisfies:

$$\mathcal{D}^Y(g)f(x) = f\big(\mathcal{D}^X(g)x\big)$$

Compositions of equivariant functions are themselves equivariant functions [**equivariant˙cohen**]. As such, we may form an equivariant network by composing it layer-wise from a set of common equivariant functions. Here, we consider four types of $SO(3)$-equivariant functions from which we may compose our equivariant networks layer-wise [**equivariant˙cohen**]. These consist of: $SO(3)$-feature convolutions, ℓ-wise self-interactions and non-linearities, and pooling across nodes. A brief description of each is given in Supplementary Note C.

In an $SO(3)$-equivariant network, features are further associated with $SO(3)$ irreducible representations: namely, spherical harmonics Y_ℓ^m, which are doubly indexed by rotational order $\ell \geq 0$, and azimuthal order $-\ell \leq m \leq \ell$. In such cases, the node-wise features $(v_{nc}^L)_\ell^m$ of channel c, are equipped with an additional two indices ℓ and m [**tensorfieldnetworks**], corresponding to the aforementioned indices of the harmonic representations.

The outputs of these networks are then also associated with some indices ℓ and m, which may be interpreted as a set of spherical harmonic basis components [**e3nn**]. These $SO(3)$ equivariant networks are naturally well suited for the prediction of tensorial properties. As will be shown below, tensors may generally be decomposed into a set of $SO(3)$ invariant subspaces, which can each be associated with a set of spherical harmonic tensors themselves. Since the outputs of $SO(3)$ networks naturally transform like spherical harmonic tensor coefficients indexed by ℓ and m, we may then essentially read them off as such and convert these into Cartesian tensor components.

3 Spherical Harmonic Tensors and $SO(3)$ Invariant Subspaces

We may construct a set of spherical harmonic tensors $\hat{Y}_\ell^m$ by first defining the J_z vector basis and then considering symmetric products of such vectors. We define the J_z basis [**mochizuki1988spherical**, **mochizuki1996spherical**, **edmonds**]:

$$\begin{bmatrix} a_+ \\ a_0 \\ a_- \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ -\frac{1}{\sqrt{2}} & +\frac{i}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

where x, y, z are Cartesian components of the same vector. In this basis, the unit vectors associated with these components correspond to unit vector spherical harmonics $\hat{Y}_{\ell=1}^m$ (where we will assume Racah normalization for all definitions[1]).

---

[1]In this work, we define the spherical harmonic functions in the Racah normalization as: $Y_\ell^m(\theta, \phi) = \sqrt{\frac{(\ell-m)!}{(\ell+m!)}} P_\ell^m(\cos\theta)e^{im\phi}$; which makes Clebsch-Gordon relations simpler.

Products of spherical harmonics then may be decomposed into sums of single spherical harmonics by way of the Clebsch-Gordon expansion [**sakurai**]:

$$Y_{\ell_1}^{m_1} \otimes Y_{\ell_2}^{m_2} = \sum_{L=-|\ell_1-\ell_2|}^{\ell_1+\ell_2} \sum_{M=-L}^{L} c_{\ell_1 0 \ell_2 0}^{L0} c_{\ell_1 m_1 \ell_2 m_2}^{LM} Y_L^M.$$

where $c_{\ell_1 m_1 \ell_2 m_2}^{LM}$ are the Clebsch-Gordon coefficients. This expansion may be used to draw a correspondence between symmetric tensor products of n vectors in the J_z basis and order $\leq n$ spherical harmonic tensors. Thus giving us a sets of basis tensors for symmetric tensor spaces that transform as irreducible representations of $SO(3)$. The coefficients for such a transformation relevant for tensors describing the material properties considered here are listed in section F.

The spherical harmonics Y_ℓ^m canonically refer to a complex basis, as should be clear from our definition of the J_z basis. Many tensors describing macroscopic responses of materials are strictly real-valued however. In the applications here we then desire to transform into a set of real-valued spherical harmonics $Y_{\ell m}$, with the transformation adopted here given in section D.

Since the basis provided by the construction method above only applies to symmetric tensor spaces, tensors of arbitrary symmetry may need to first be decomposed into a set of totally symmetric subspaces. We may decompose an arbitrary tensor of order-n into a set of symmetric tensors of order-ℓ with $0 \leq \ell \leq n$ by projecting it onto it's irreducible $SO(3)$-invariant subspaces. We then may use the relations between the J_z basis components and the real spherical harmonic components $y_{\ell m}$ using the relations discussed above.

$SO(3)$ invariant subspaces can be constructed via a compound decomposition with respect to GL, SL, and O: which result from application of Young symmetrizers, contractions with the totally antisymmetric ϵ, and contractions with the totally symmetric g, respectively. For an overview of these decompositions, see section A.

As a basic example, consider a general order-two tensor, as displayed in fig. 1, where we may first construct the totally symmetric and antisymmetric subspaces, and then contract with the metric tensor g_{ij} and the totally antisymmetric Levi-Civita tensor ϵ_{ij}, respectively. This achieves an irreducible $SO(3)$ decomposition of the space into a scalar (order-zero) trace s, a symmetric order-two residue R, and a scalar antisymmetric component a.

In the manner described above, we are able to decompose an arbitrary tensor into a set of components that transform like spherical harmonics, just like the representations passed throughout, and output by, $SO(3)$ equivariant models. To this end, we now give spherical harmonic decompositions of several tensors describing common material properties, and then use these decompositions to predict these tensors directly from material structure.

4   Spherical Harmonic Decomposition of Common Material Property Tensors

Here is presented a brief overview of the harmonic decomposition of three common, tensor-valued material properties: the dielectric tensor ϵ, the piezoelectric strain tensor d, and the elasticity tensor C. A more systematic presentation of
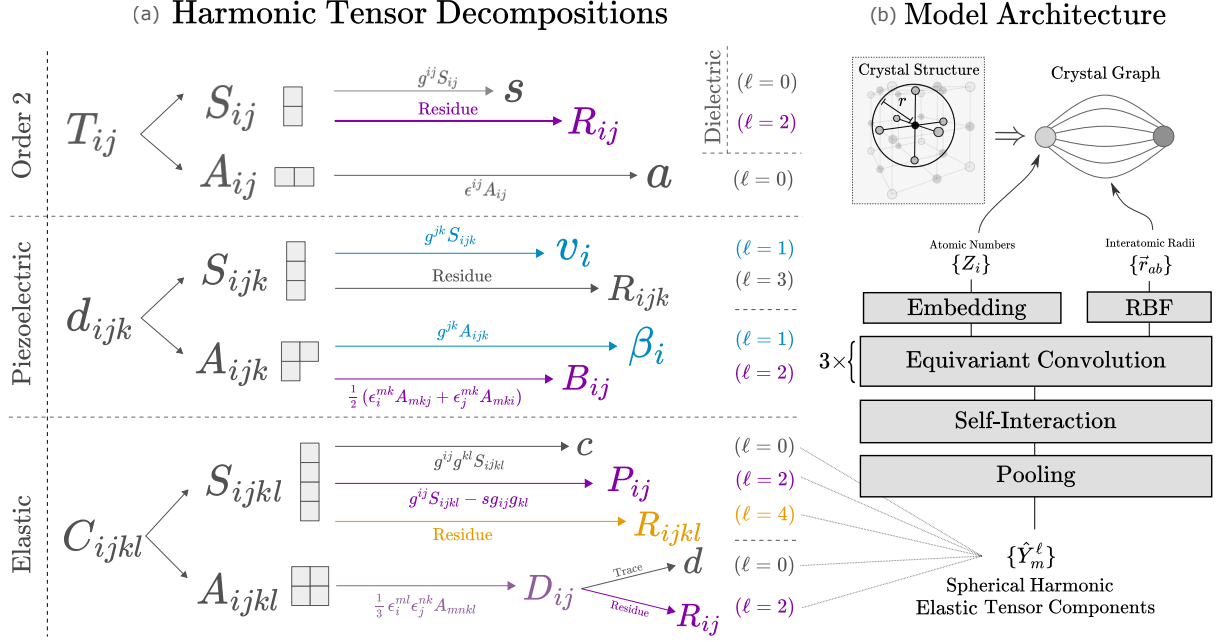
Figure 1: Overview of $SO(3)$ decomposition of order two tensors, piezoelectric (order 3) tensors, and elasticity (order 4) tensors. Grey corresponds to scalar subspaces of the related tensor space, blue to order 1 subspaces, purple to order 2, and yellow to order 4. A schematic for the general model architecture used in this work is also provided, where the output coefficients correspond directly to harmonic tensor components resulting from these decompositions.

each decomposition is given in reference section B. Note that in the following, we ignore the distinction between co- and contravariant components due to the Euclidean structure $g_{ij} = \delta_{ij}$.

Dielectric Tensors

The dielectric permittivity tensor ϵ of some material is a linear model of it's electric displacement $\vec{D}$ in response to an external electric field $\vec{E}$:

$$\vec{D} = \epsilon \vec{E}$$

Here we focus only on static dielectric responses, that is, $(\partial/\partial t)\vec{E} = 0$, primarily due to the restriction of data to this case. The dielectric tensor ϵ is an order-two tensor formed of vector spaces over the field $\mathbb{R}^3$. The dielectric tensor ϵ is symmetric under permutation of it's indices, such that:

$$\epsilon_{ij} = \epsilon_{ji}.$$

The dielectric tensor then may be decomposed further into it's trace $s = g_{ij}\epsilon_{ij}$ and it's traceless residue R, defined by:

$$R_{ij} = \epsilon_{ij} - sg_{ij}$$

5

resulting in two invariant subspaces of order-0 and 2, respectively. These two symmetric subspaces then admit a decomposition in the manner described in section 3.

Piezoelectric Tensors

The piezoelectric strain constants $(d_{ijk})_T$ are defined (at constant temperature) by the thermodynamic relation:

$$(d_{ijk})_T = \left(\frac{\partial \epsilon_{ij}}{\partial E_k}\right)_{\sigma,T}$$

where here ϵ is the strain tensor, and the partial derivative is taken at constant stress σ and temperature T. These strain constants d_{ijk} are related to the piezoelectric stress constants e_{ijk} via the elastic tensor C_{ijkl} according to:

$$(e_{ijk})_T = (d_{ilm})(C_{ijkl})_{E,T}$$

with e_{ijk} defined as:

$$(e_{ijk})_T = \left(\frac{\partial D_i}{\partial \epsilon_{jk}}\right)_{\sigma,T}$$

and where D is the resulting electric displacement vector in the material (we now generally neglect the explicit notation of constant parameters). The piezo-electric strain components d_{ijk} are symmetric under i,j due to the symmetry of the strain tensor ϵ_{ij}, so that we have:

$$d_{ijk} = d_{jik}.$$

This symmetry gives us a natural GL decomposition into the totally symmetric S and the mixed-symmetry A, defined component-wise from d as [itin-rank3]:

$$S_{ijk} = \frac{1}{3}\big(d_{ijk} + d_{ikj} + d_{kji}\big)$$
$$A_{ijk} = \frac{1}{3}\big(2d_{ijk} - d_{ikj} - d_{kji}\big).$$

The fully symmetric part S then consists of a trace vector $s_k = g_{ij}S_{ijk}$, and a symmetric residue W_{ijk} defined by:

$$W_{ijk} = S_{ijk} - \frac{1}{5}(g_{ij}v_k + g_{ik}v_j + g_{jk}v_i)$$

These correspond to the spaces $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(3)}$, respectively. The mixed symmetry part A also requires decomposition with respect to $SO(3)$, so that we have a set of symmetric tensors describing it. It's 8 independent components can be described by a $5 \oplus 3$ dimensional space consisting of a symmetric order-2 tensor and a trace vector.

The trace vector a^i, describing A's 3-dimensional $SO(3)$ invariant subspace, can be formed from the contraction of the metric tensor g_{ij} along A's first and second indices. That is, we define:

$$a^i = g_{jk}A^{ijk}$$

Note that this choice is somewhat arbitrary, since we could define the trace part to correspond to the contraction along the first and third indices, or the second and third. However, it can be shown that for the mixed symmetry of A, these two trace vectors are linearly dependent (related by an overall factor of 1 and -2, respectively).

The order-2 invariant subspace of A then may be constructed by symmetrizing the partial contraction with ϵ along the first and third indices (the antisymmetric pair). Note that the antisymmetric part of this partial contraction corresponds to the trace vector space accounted for here by u. Explicitly, we define:

$$b_{ij} = \frac{1}{2}\big(\epsilon_{nim}A_{njm} + \epsilon_{njm}A_{nim}\big)$$

which is a traceless symmetric order-2 tensor.

And so, as mentioned above, the symmetric part S is readily decomposed in the spherical bases into $\mathcal{H}^{(1)} \oplus \mathcal{H}^{(3)}$. The mixed-symmetry part then inhabits the space $\mathcal{H}^{(1)} \oplus \mathcal{H}^{(2)}$.

Elasticity Tensor

The elasticity (or simply, elastic) tensor C of a material relates it's Cauchy strain ϵ to some infinitesimal stress τ [**landau-elastic**] . This may be described, component-wise, as the linear relation below:

$$\tau_{ij} = C_{ijkl}\epsilon_{kl}$$

As should be clear by the number of indices, the elastic tensor is a fourth order tensor.

The elastic tensor has several symmetries, but is not totally symmetric in general. For all material systems, C must satisfy the so-called 'minor symmetries' below:

$$C_{ijkl} = C_{jikl}$$
$$C_{ijkl} = C_{ijlk}$$

resulting from the symmetry of the strain and stress tensors ($\tau_{ij} = \tau_{ji}$ and $\epsilon_{ij} = \epsilon_{ji}$) under the assumption of equilibrium. This reduces the number of independent components from 81 to 36.

For conservative systems in which the elastic deformation is describable in terms of some potential energy function (and which we shall henceforth assume for our application), C has the additional 'major symmetry':

$$C_{ijkl} = C_{klij}$$

This further reduces the number of independent components from 36 (from the minor symmetries) to 21 in total.

In a GL-decomposition, these symmetries leave us with two non-vanishing subspaces: the totally symmetric S, and a mixed symmetry A (section B.3), defined component-wise as [**itin-elastic**]:

$$S_{ijkl} = \frac{1}{3}\big(C_{ijkl} + C_{ikjl} + C_{klij}\big)$$

$$A_{ijkl} = \frac{1}{3}\big(2C_{ijkl} - C_{ikjl} - C_{klij}\big)$$

(this mixed-symmetry subspace A also corresponds to Backus' [**backus1970geometrical**] asymmetric tensor A). The fully symmetric part can of course be converted to spherical harmonic components via a Clebsch-Gordon expansion. The mixed symmetry A, however requires further decomposition.

All six independent components of A can be described by the symmetric (but not traceless) tensor t, defined as the double partial contraction of A with the totally antisymmetric tensor ϵ as:

$$t_{ij} = \epsilon_i^{mk} \epsilon_j^{nl} A_{mnkl}$$

This tensor t then has a harmonic decomposition according to the order-two Clebsch-Gordon transformation between the J_z basis and the harmonic basis y_ℓ^m.

As such, the elastic tensor C admits a decomposition into S and A, which inhabit spaces of $\mathcal{H}^{(4)} \oplus \mathcal{H}^{(2)} \oplus \mathcal{H}^{(0)}$ and $\mathcal{H}^{(2)} \oplus \mathcal{H}^{(0)}$, respectively.

5   Methods

Results presented here investigate the performance of three distinct $SO(3)$ equivariant models in the direct prediction of tensor components from crystal structure. This is achieved in the manner outlined above, wherein we utilize equivariant model's output as a set of spherical harmonic coefficients and then convert these to Cartesian tensor components.

Model Architecture

The model architecture used in this work is based on the design of the SE(3) Equivariant graph neural networks found in **brandstetter2021geometric**. This architecture consists of an layer embedding the input atom features into spherical harmonics up to order ℓ_{max}. These features are then passed through R graph convolution layers (either message passing, convolution, or transformer layers) where nodes are updated by their neighbors. The output representation is then passed through N self interaction layers in which individual nodes update themselves. A pooling layer (either mean or max pooling) is then applied across node indices to return a crystal-level feature vector, which is then passed through M MLP layers. This output, still associated with ℓ and m indices, is then used to predict the coefficients of target tensors. A schematic for this architecture is provided in fig. 1.

Three equivariant convolutional layers were tested and compared in this model architecture: the Steerable Equivariant Graph Neural Network (SEGNN) and the Steerable Equivariant Non-Linear Convolutional layer (SEConv) [**brandstetter2021geometric**]; and the Steerable Equivariant Transformer (SETransformer) [**fuchs2020setransformer**]. All are based on the popular e3nn python module [**e3nn˙paper**], with features and outputs generally associated with a set of spherical harmonic indices. Note that the SETransformer is unique in that it alone uses attention in the filter function $\mathcal{F}_{acm_im_o}^{\ell_o \ell_i}(r_{ab}, v_a, v_b)$, such that edge weights may be learned from origin and connected node features v.

Data

The data utilized is available in the *Materials Project* database [**matproj**] which includes ∼150k materials. Of which, almost all have calculated scalar values for band gap and formation energy, ∼12k have a calculated elastic tensor [**elastic-matproj**], ∼6k have a calculated dielectric tensor [**di-matproj**], and ∼3k have a calculated piezoelectric tensor [**piezo-matproj**]. The data was split 70% for training and 15% for validation, with a remaining 15% held out for testing. Data distribution statistics can be found in section G.

Graphs input into the model were derived from the material structures available in the database. For up to the first twelve neighbors of atoms within 8 Angstrom, edges were constructed. These edges' features consisted of spherical harmonics of orders $0 \leq \ell \leq \ell_{max}$ evaluated on each edge's corresponding unit vector direction, with the distance between atoms being included as $\ell = 0$ features. Initial node features then encoded the following atomic information: atomic number[2], group[2], period[2], block[2], electronegativity, atomic radius, number of valence electrons, electron affinity, ionization energy, and atomic mass; following the method adopted in [**cgcnn**].

6   Results

Here we present results for the prediction of tensor components from a model of the form discussed above, and compare the relative performance of three distinct equivariant convolutional layers in section 6.1.

One significant hurdle in the prediction of tensorial material properties is the scarcity of training data. Pretraining on a diverse set of related targets may be particularly useful in this case [**pretraining**]. To test the transferability of tensor data to different tasks, we train convolutional layers through a variety of target set progressions. Models are first pretrained on one of two scalar datasets: band gap or formation energy, and then compared to un-pretrained benchmarks in section 6.2. These are then compared to cascades of pretraining on different tensorial target sets in section 6.3. Finally, a set of common scalar elastic values are calculated from the predicted elasticity tensors, with such predictions discussed in section 6.4.

6.1   Tensor Prediction

Three distinct models were trained to predict tensor targets component-wise from a set of randomly initialized parameters, and then tested on a set of test data withheld through training. Test results for these predictions are given in table 1, where metric used to report results for prediction tasks is average MAE over components, following the choice of metric for such predictions in [**StrainNet**]. Component-wise MAE is presented in fig. 2.

The average MAE across tensor components in testing beats previous results as reported in [**StrainNet**], with an improvement from 10.37 GPa to 7.69 GPa with SEConv models in this work. Best results for other tensorial targets were 4.7 with the SEConv model, and 0.17 C/m^2 with both SEGNN and SEConv.

---

[2] These properties were one-hot encoded.

Since some components must be zero for high-symmetry systems, these components may be unfairly influential in a naive scalar metric based on an average across all components for all systems. Indeed, heatmaps displaying component-wise MAE in section 6.1 suggests incorporation of symmetry class information is necessary to further improve results, as in [**yan2024equivcrystalprediction**]. This should be apparent in the lower MAE contributions in several off-diagonal components in the dielectric and elasticity prediction heatmaps, where in an average over all systems, many must be zero due to symmetry constraints.
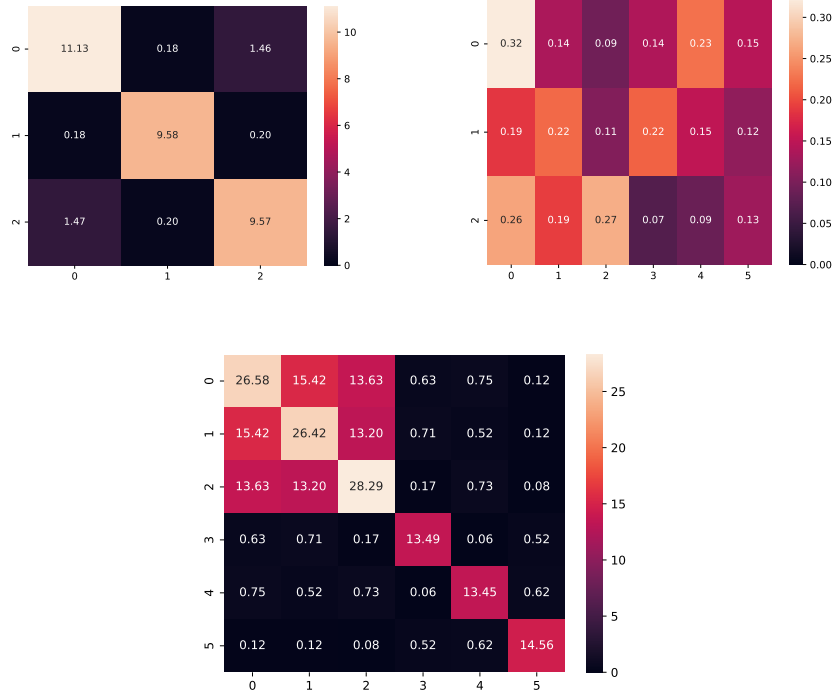
Figure 2: Heatmaps showing the component-wise MAE in Voigt notation for the all direct prediction tasks. Shown in order of dielectric, piezoelectric, and elastic tensor components, respectively. These prediction statistics are from a model with SEGNN convolutional layers.

6.2 Scalar Pretraining & Transfer Learning Applications

Transfer-learning experiments here focus on a filter-down approach, in which we start training on larger datasets and then transfer the trained graph network weights to a downstream task with a smaller dataset (while continuing to train the transferred network).

Since all tensorial data is relatively scarce, pretraining on a scalar target set would be ideal, since such data is the most plentiful for material systems. To test such applications, we here consider training progressions of the following

Target Components	MAE (Averaged Over Components)		
	SEGNN	SEConv	SETransformer
Elastic *[GPa]*	8.139	7.689	7.941
Dielectric	4.82	4.702	4.718
Piezoelectric *[C/m²]*	0.170	0.170	0.1714

Table 1: Results for direct prediction of elastic, piezoelectric, and dielectric tensor components.

form:

$$Scalar\ Target \to Tensor\ Target,$$

The two scalar datasets used in pretraining included the predicted formation energy per atom for a crystal, and the calculated band gap.

Results from these experiments can be found in table 2. Note that in these progressions, no data was witheld in the scalar target pretraining process. In general, pretraining of this nature had little effect on down-stream results. For predictions after pretraining on the scalar formation energy, the prediction accuracy across components only improved substantially for the SETransformer model: with average MAE decreasing to 7.65 GPa (from 7.94 GPa) for elasticity components, and 4.41 (from 4.72) for dielectric targets. Pretraining on band gap had even less effect on down-stream performance, except for a notable improvement in the SEConv model's dielectric average test MAE from 4.7 to 4.35. Other predictions had little-to-no improvement after scalar pretraining.

Poor-to-no improvement after such pretraining may be due to a number of factors: perhaps there is little similarity between pretraining and end target datasets, so that pretraining overgeneralizes the model to a larger class of materials than are relevant. Alternatively, the difference in the rotational order of outputs between tasks may also result in little overlap between the pretrained filter weights and the weights relative to the down-stream task. This suggests pretraining on different tensor targets of similar order may improve down-stream task performance more than scalar pretraining sets.

Poor transfer-learning results may also be due to little domain relevance between different targets. This is a point reinforced by the most substantial improvements being observed in arguably more related target set pairs: namely, from formation energy to elasticity predictions in the SETransformer; and the pretraining progression of band gap to dielectric predictions in the SEConv progression.

6.3 Chained Pretraining for Piezoelectric Tensors

The smallest dataset considered here is that of piezoelectric strain constants. Performance did not significantly improve with any scalar target pretraining in any of the models for this task. The cost and scarcity of data for this target motivates further search for improvement from other tensorial target sets. To this end, the experiments performed here focus on three different training progressions of tasks:

1. Formation energy/Band gap $\to$ Elastic $\to$ Dielectric $\to$ Piezoelectric

2. Formation energy/Band gap $\to$ Elastic $\to$ Piezoelectric

Pretraining Task:	Tensor Target	MAE after pretraining (*without pretraining*)		
		SEGNN	SEConv	SETransformer
Formation Energy	Elastic *[GPa]*	8.19 *(8.14)*	7.98 *(7.69)*	7.65 *(7.94)*
	Dielectric	5.02 *(4.82)*	4.52 *(4.70)*	4.41 *(4.72)*
	Piezoelectric *[C/m^2]*	0.170 *(0.170)*	0.170 *(0.170)*	0.171 *(0.171)*
Band Gap	Elastic *[GPa]*	8.39 *(8.14)*	7.87 *(7.69)*	8.34 *(7.94)*
	Dielectric	4.74 *(4.82)*	4.35 *(4.70)*	4.56 *(4.72)*
	Piezoelectric *[C/m^2]*	0.170 *(0.170)*	0.170 *(0.170)*	0.171 *(0.171)*

Table 2: Results for tensor component prediction accuracy after pretraining on a scalar-target task for materials. Numbers italicized in parentheses are the original accuracies (without pretraining).

Pretraining Task:	MAE (Avg. Over Components)		
	SEGNN	SEConv	SETransformer
Formation Energy	4.67 *(4.82)*	4.69 *(4.71)*	4.45 *(4.72)*
Band Gap	4.66 *(4.82)*	4.64 *(4.71)*	4.74 *(4.72)*

Table 3: Results for predicting the components of the dielectric tensor in experiment 1.

3. Formation energy/Band gap → Dielectric → Piezoelectric

Overall, there was no change in average prediction accuracy for the piezoelectric tensor through all experiments of the above form, with an average MAE remaining 0.17 C/m^2 for all. The reasons for this lack of improvement may be similar to those given in the scalar pretraining discussion: disjointed target data, and/or filter weights due to differing rotational orders between tasks. However, there should be atleast some overlap in filter weight training due to the common prediction of order $\ell = 2$ outputs between all three tensor targets. It may be that the scarcity of piezoelectric target data alone is the fundamental restriction on performance, in which case a larger set of down-stream training data would improve the effectiveness of such pretraining strategies.

6.4 Scalar Elastic Properties

The bulk modulus K, shear modulus G, and Young modulus E, are common scalar-valued quantities representing elastic responses in materials. These values are computed from the elasticity tensor C and compliance tensor $s = C^{-1}$ as the average of the Voigt and Reuss forms as in section E. The model used in this work performs worse than MatTen [**MatTen**] on these scalar tasks, with best results for these scalar predictions including an MAE of 8.41 GPa (compared to 7.37 GPa in [**MatTen**]) on the K_H test set with the SEConv model, and then 9.04 GPa on the G_H set (compared to 8.38 GPa) and 20.06 GPa on E (compared to 20.59 GPa) with SEGNN.

Performance in the prediction of these components is presented in section 6.4, where pretraining on scalar target sets is also considered. Pretraining on band gap shows no benefit. However, the pretraining task of formation energy did provide slight improvements for the SEConv and SETransformer models. The most notable improvements were observed with the SETransformer model, with

Model	Pretraining	MAE		
		K_H (GPa)	G_H (GPa)	E (GPa)
SEGNN	None	8.61	9.04	20.06
	Band Gap	9.08	9.94	22.18
	Formation Energy	8.90	9.06	20.16
SEConv	None	8.41	9.37	21.61
	Band Gap	8.46	9.30	20.80
	Formation Energy	8.42	8.81	19.56
SETransformer	None	9.67	9.51	21.24
	Band Gap	10.66	10.18	23.04
	Formation Energy	10.06	8.99	20.35
MatTen [**MatTen**]	None	7.37	8.38	20.59

Table 4: The MAE of the prediction of the bulk modulus, K, shear modulus, G, and the young modulus, E, derived from the predicted elastic tensor in comparison to MatTen [**MatTen**].

test MAE improving from 21.14 GPa to 20.35 GPa on the Young modulus dataset, and an improvement from 9.51 GPa to 8.99 GPa on the shear modulus dataset.

7   Conclusion

As outlined in this work, $SO(3)$ equivariant networks may be used to directly predict tensor components in a manner consistent with differing choices of basis for input features due to the equivariance of the network. Results presented here focus on three common tensorial properties of solid state materials describing dielectric, piezoelectric and elastic responses. However, the approach taken here may be applied to arbitrary tensorial targets by first decomposing into their GL-invariant subspaces by way of Young symmetrizers, then their $SO(3)$-invariant subspaces and finally transforming these symmetric tensor subspaces into a spherical tensor basis.

In a certain sense, the component-wise results and lack of transferability may indicate that the spherical harmonic approach is potentially 'too big' or too expressive to effectively predict crystalline material properties in a way that naturally respects crystalline symmetries. Indeed, $SO(3)$ convolution, as described here, generates a large set of unwanted and consequently discarded combinations of spherical harmonics. Future works may incorporate the symmetry of these crystalline systems as in [**yan2024equivcrystalprediction**], so that models may naturally predict which tensor components should be zero due to symmetry alone. A simple approach would be to mask results according to crystal system, or build an individual model for each different crystal system with a unique output structure. A more sophisticated approach might be to restrict convolution specifically to the crystal group elements, as opposed to the larger $SO(3)$ group.

Code Availability

The code used in this work can be found in the following GitHub repository:
https://github.com/qmatyanlab/Transfer-Learning-For-Complex-Crystal-Properties.

A Constructing $SO(3)$-Invariant Subspaces of Tensors

An invariant subspace of a tensor is a subspace which is closed under the action of some group of transformations. That is, invariant and disjoint subspaces of a tensor shouldn't 'mix' at all under the corresponding group transformations.

Invariants under $SO(3)$ of an arbitrary tensor T may be constructed by way of Young symmetrizers, and subsequent contractions with the metric tensor g_{ij} or the fully antisymmetric tensor ϵ_{ijk}. Note that these three methods correspond to the decomposition of a tensor with respect to the general linear group (by way of Young symmetrizers), the orthogonal group (contractions with g_{ij}), and the special linear group (contractions with ϵ_{ijk}), respectively.

Since $SO(3) = SL(3) \bigcap O(3) \subset GL(3)$, to construct invariants of SO, we generally form invariant subspaces under GL by way of Young symmetrizers first, and then construct invariants under O and/or SL of these GL invariant subspaces. This often results in several different harmonic subspaces of the same order ℓ in the harmonic decomposition of a tensor. Also, note that the GL decompositions for tensors of order greater than two are not, in general, unique.

Below, we introduce the tools used in the construction of such invariant subspaces and demonstrate their invariance. These tools are then applied in three cases of tensor spaces with importance in the field of materials science: corresponding to the space of dielectric tensors, piezoelectric response tensors, and elasticity tensors.

A.1 GL Decompositions

By way of the Schur-Weyl duality, the invariant subspaces of a tensor under the general linear group GL determine the invariant subspaces of the tensor under the symmetric group S_n, where symmetric group elements act as permutations on tensor indices.

Under the symmetric group, the invariant subspaces of a tensor are described by way of standard Young tableaux, and then constructed from an arbitrary tensor by means of the corresponding Young symmetrizers. Below, these concepts are briefly introduced for the purpose of GL decompositions of tensors over $\mathbb{C}^n$. Note that these decompositions are in general not unique for tensors of order > 2, so some conventions must be adopted. In the present work we adopt the convention of antisymmetrization before symmetrization in the construction of Young symmetrizers.

A.1.1 Young Diagrams, Tableaux and Symmetrizers

Young diagrams of order n are left-justified arrangements of boxes into k rows stacked vertically in non-increasing order. A Young diagram is said to be of some shape $\lambda : (\lambda_1, \lambda_2, ..., \lambda_k)$, where λ_i refers to the depth of row i and $\lambda_{i+1} \leq \lambda_i \leq \lambda_{i-1}$.

We can form a set of Young tableaux from diagrams by filling in the boxes with indices from an ordered set $\{x_1, x_2, ..., x_k\}$ corresponding to tensor components as $T^{x_1 x_2 \cdots x_k}$. A standard tableu is one filled with indices x_i (without
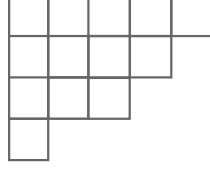
Figure 3: Example Young diagram of shape $(5, 4, 3, 1)$.
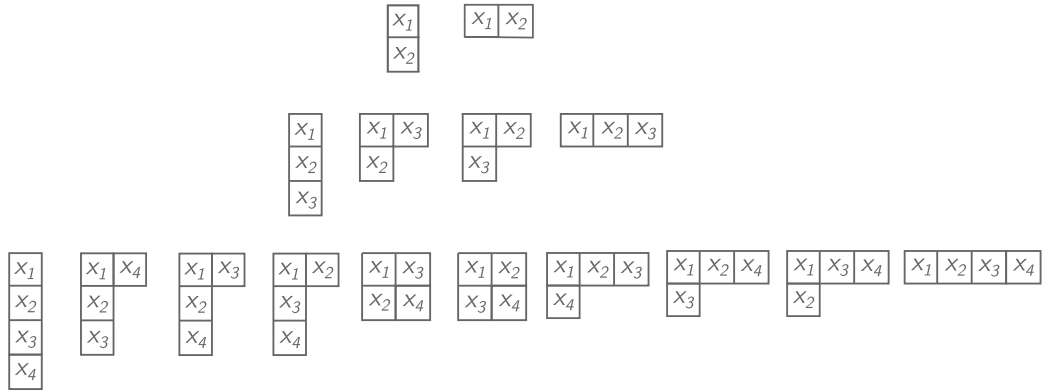
Figure 4: Standard Young tableaux for diagrams of at most 4 boxes.

repeats) with entries increasing in index i down each column and across (to the right) rows.

Each of these standard tableaux correspond to an invariant subspace under S_k, and further GL_n, by way of the Schur-Weyl Duality. From these tableau, we may construct so-called *Young Symmetrizers*, which project tensors onto their corresponding S_k-invariant subspace.

A Young symmetrizer P_λ corresponding to a tableau λ is composed of a compound set of symmetrizing operations s_λ and antisymmetrizing operations a_λ, scaled by an overall normalization constant C_λ:

$$P_\lambda = \mathcal{C}_\lambda s_\lambda a_\lambda$$

Where here, we adopt the convention of anti-symmetrization before symmetrization, following that in Itin [**itin-rank3**].

The symmetrizing operator s_λ for a diagram λ is composed of a product of symmetrizers $\mathcal{S}(\mathcal{I})$, where $\mathcal{I}$ ranges over all subsets of indices corresponding to some vertically stacked set of indices in tableau λ, i.e.:

$$s_\lambda = \prod_{\mathcal{I} \in \mathrm{Cols}(\lambda)} \mathcal{S}(\mathcal{I})$$

where $\mathrm{Cols}(\lambda)$ represents the set of disjoint subsets of indices down each column,

and symmetrizers $\mathcal{S}$ are defined to act on tensors T component-wise as:

$$\big[\mathcal{S}(\mathcal{I})T\big]_{ijk\ldots} = \sum_{\sigma_{\mathcal{I}}} T_{\sigma_{\mathcal{I}}(ijk\ldots)}$$

where $\sigma_{\mathcal{I}}$ are permutations of index subset $\mathcal{I}$. Similarly, the antisymmetrizing operator a_λ can be constructed as a product of antisymmetrizers:

$$a_\lambda = \prod_{\mathcal{I} \in \mathrm{Rows}(\lambda)} \mathcal{A}(\mathcal{I})$$

where $\mathrm{Rows}(\lambda)$ represents the set of disjoint subsets of indices across each entire row, and antisymmetrizers $\mathcal{A}$ are defined as:

$$\big[\mathcal{A}(\mathcal{I})T\big]_{ijk\ldots} = \sum_{\sigma_{\mathcal{I}}} \mathrm{sgn}(\sigma_{\mathcal{I}}) T_{\sigma_{\mathcal{I}}(ijk\ldots)}$$

where, again, $\sigma_{\mathcal{I}}$ range over all permutations of index subset $\mathcal{I}$.

The normalization constant C_λ may be derived from the shape of the underlying Young diagram according to the *hook-length formula*, given below, where $\mathrm{hook}(\alpha, \beta)$ returns the number of boxes crossed by a hook coming up (from below) column β and out of the diagram to the right in row α.

$$C_\lambda = \prod_{(\alpha,\beta)\in\lambda} \frac{1}{\mathrm{hook}(\alpha, \beta)}$$

Furthermore, the number of independent components N_λ of a tensor subspace corresponding to some Young tableau with shape λ may also be derived from the diagram by a related hook-length formula:

$$N_\lambda = \prod_{(\alpha,\beta)\in\lambda} \frac{n - \alpha + \beta}{\mathrm{hook}(\alpha, \beta)}$$

where n is the dimension of the vector space forming the tensor space.

Thus, the standard Young tableaux of k boxes can be used to decompose an arbitrary tensor space into a set of GL invariant subspaces with known symmetries (under permutation of indices) by way of corresponding Young symmetrizers. These known symmetries will be relevant in the further decomposition of these subspaces by way of contractions with the metric tensor, and the fully antisymmetric tensor, discussed below.

A.2 SL Decompositions

The special linear group SL of transformations is defined as the subset of invertible linear transformations with determinant equal to positive one. Under SL, orientation and volume are preserved, where volume is defined as the contraction of a tensor with the fully antisymmetric tensor ϵ_{ijk}, which transforms under $R \in GL(3)$ as:

$$\epsilon_{x_1 x_2 x_3} \to \epsilon_{x_1' x_2' x_3'} = \det(R) R^{x_1}_{x_1'} R^{x_2}_{x_2'} R^{x_3}_{x_3'} \epsilon_{x_1 x_2 x_3}$$
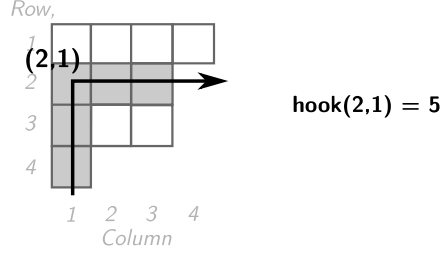
Figure 5: Example value of hook(2, 1) for a diagram of shape $\lambda : (4, 3, 3, 1)$. Note that it's corresponding normalization constant is $C_\lambda = 1/33600$.

For $R \in SL$ then, we have $\det(R) = 1$, allowing us to form various invariants of tensors by way of contraction (or partial contraction) with ϵ. For example, for a order-three tensor T, the volume may be defined component-wise as:

$$V = \epsilon_{ijk} T^{ijk}$$

This V is then clearly invariant under SO since for any transformation $R \in SO(3)$, we have:

$$V \to \epsilon_{i'j'k'} T^{i'j'k'} = R^i_{i'} R^j_{j'} R^k_{k'} \epsilon_{ijk} R^{i'}_i R^{j'}_j R^{k'}_k T^{ijk} = \epsilon_{ijk} T^{ijk}$$

Note that contractions with ϵ along fully symmetric sets of indices will always vanish; a fact useful when considering particular GL invariant subspaces. Explicitly,

$$\epsilon_{ijk} T_{..(ijk)..} = 0$$

A.3 O Decompositions

The orthogonal group O of transformations is defined as the subset of invertible linear transformations satisfying $R^T R = 1$. Under O, we may define an inner product between vectors $< \cdot, \cdot >$ by means of a metric tensor g_{ij} (δ_{ij} in Euclidean space) , which transforms as a second order tensor under some transformation R as:

$$g_{x_1 x_2} \to g_{x'_1 x'_2} = R^{x_1}_{x'_1} R^{x_2}_{x'_2} g_{x_1 x_2}$$

The inner product of vectors u, v is defined component-wise as:

$$< u, v > = u^i g_{ij} v^j$$

which is then clearly invariant under a transformation $R \in O(3)$ as:

$$< u', v' > = R^{i'}_i u^i R^j_{j'} R^j_{j'} g_{ij} R^{j'}_j v^j = u^i g_{ij} v^j = < u, v >$$

Furthermore, we may construct different invariants for different order tensor spaces. For example, we may define O invariant scalar-valued traces $Tr(\cdot)$ of second order tensors M by means of the metric tensor, defined component wise as

$$Tr(M) = g_{ij} M^{ij}$$

and we may form three O invariant trace vectors for third order tensors T, defined component wise as:

$$\nu^k = g_{ij}T^{ijk}, \quad \mu^j = g_{ki}T^{ijk}, \quad u^i = g_{jk}T^{ijk}$$

That is, with the metric tensor g_{ij} available, as is the case in $SO(3) \subset O(3)$, we may decompose a order-n tensor into a sets of order-$(n-2), (n-4), ..., (0 \text{ or } 1)$ tensors by taking sequential traces.

The trace-part of a tensor may then be reconstructed by taking tensor products of metric tensors and trace vectors. A point which is relevant to GL invariant subspaces is that contractions with respect to g under fully antisymmetric pairs of indices will always vanish. That is,

$$g_{ij}T_{..[ij]..} = 0$$

SO(3) Decompositions

In $SO(3)$, we have all of the above tools (Young diagrams, g_{ij}, and ϵ_{ijk}) available. As such, we may decompose a tensor into a set of $SO(3)$ invariant subspaces by their use, and all will be valid. Unfortunately, for tensors of order> 2, there is not, in general, a unique and irreducible set of invariant $SO(3)$ spaces. However, we may always decompose an arbitrary tensor into a set of (not necessarily unique) irreducible $SO(3)$-invariant symmetric tensor subspaces using the tools given above.

B   Full Decompositions of Tensors

Below, we give the $SO(3)$ decomposition of the three tensors of interest here. Namely, the order-two dielectric tensor ϵ, the order-three piezoelecrtic strain tensor d, and the order-four elastic tensor C. Each begins with the Young diagrams for the total order tensor space and then uses the symmetries of the respective tensor to restrict down to the relevant Young tableaux. These are used to form symmetrizers and then further decomposed with contractions with the metric tensor g and the levi-civita tensor ϵ until all that remains are a set of fully symmetric sub-tensors.

B.1   Dielectric Tensor Decomposition

The decomposition of the dielectric tensor is essentially trivial since it is already symmetric. However, we give a detailed overview here for instructive purposes.

The GL decomposition of a order-two tensor space is the usual symmetric-antisymmetric decomposition of a matrix M.

$$M_{ij} = S_{ij} + A_{ij}$$

with:

$$S_{ij} = \frac{1}{2}(M_{ij} + M_{ji}) \quad \Rightarrow \quad S_{ij} = S_{ji}$$

$$A_{ij} = \frac{1}{2}(M_{ij} - M_{ji}) \quad \Rightarrow \quad A_{ij} = -A_{ji}$$

which correspond to the Young diagrams below:

.
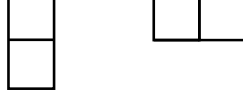
Since the dielectric tensor ϵ is symmetric, we are left then only with the unique tableaux below, corresponding to the fully symmetric part S.

We can then further decompose this order-two symmetric into a order-zero space and a order-two space by taking the order-0 trace $t = g_{ij}\epsilon_{ij}$ and then forming the order-2 traceless residue $R_{ij} = \epsilon_{ij} - t\delta_{ij}$.

B.2   Piezoelectric Tensor Decomposition

The Young diagrams for a order-three tensor are as follows:
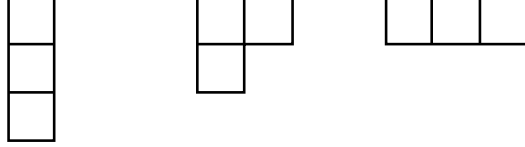
according to the symmetry $d_{ijk} = d_{ikj}$, we can again see that all Young tableaux but the following will disappear, since the rest are asymmetric in the last two indices.
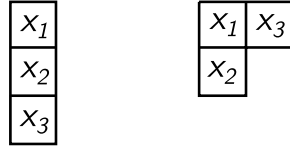
where the totally symmetric component tensor here we define as S and the mixed symmetry component tensor we define A, so that we have the Young decomposition:

$$d_{ijk} = S_{ijk} + A_{ijk}$$

defined component-wise in terms of strain components:

$$S_{ijk} = \frac{1}{3}\left(d_{ijk} + d_{ikj} + d_{kji}\right)$$

$$A_{ijk} = \frac{1}{3}\left(2d_{ijk} - d_{ikj} - d_{kji}\right)$$

The fully symmetric part S is of an adequate form for harmonic decomposition, with those relations given in the next section.

The mixed symmetry part A however requires further decomposition with respect to $SO(3)$, so that we have a set of symmetric tensors describing it. It's 8 independent components can be described by a $5 \oplus 3$ dimensional space consisting of a symmetric order-2 tensor and a trace vector.

The trace vector v^i, which corresponds to A's 3-dimensional $SO(3)$ invariant subspace, can be formed from the contraction of the metric tensor g along A's first and second indices. That is, we define:

$$v^i = g_{jk} A^{ijk}$$

Note that this choice is somewhat arbitrary, since we could define the trace part to correspond to the contraction along the first and third indices, or the second and third. However, it can be shown that for the mixed symmetry of A, these two potential trace vectors are linearly dependent (related by an overall factor of 1 and -2, respectively).

Rather simply then, we can reconstruct a third-order tensor V, corresponding to this order-one invariant subspace, by defining V component-wise as:

$$V_{ijk} = \frac{1}{4} \Big[v_i g_{jk} + v_j g_{ik} - 2 v_j g_{ik} \Big]$$

The order-2 invariant subspace of A then may be constructed by symmetrizing the partial contraction with ϵ along the first and second indices (the anti-symmetric pair). Note that the antisymmetric part of this partial contraction corresponds to the trace vector space accounted for here by u. Explicitly, we define:

$$b_{ij} = \frac{1}{2} \big(\epsilon_i^{mk} A_{mkj} + \epsilon_j^{mk} A_{mki} \big)$$

which is a traceless symmetric order-2 tensor. And from which we may reconstruct the corresponding order-3 tensor B, defined by:

$$B_{ijk} = \frac{1}{3} \Big[\epsilon_{ik}^p b_{pj} + \epsilon_{ij}^p b_{pk} \Big]$$

Thus, we may provide a total harmonic decomposition for A in terms of the invariant subspaces of the order-1 u and the order-2 b, such that:

$$A_{ijk} = V_{ijk} + B_{ijk}$$

And so, as mentioned above, the symmetric part S is readily decomposed in the spherical bases into $\mathcal{H}^{(1)} \oplus \mathcal{H}^{(3)}$. And then the mixed-symmetry part inhabits the space $\mathcal{H}^{(1)} \oplus \mathcal{H}^{(2)}$.

B.3   Elastic Tensor Decomposition

If we examine the Young diagrams of a fourth order tensor, as below:
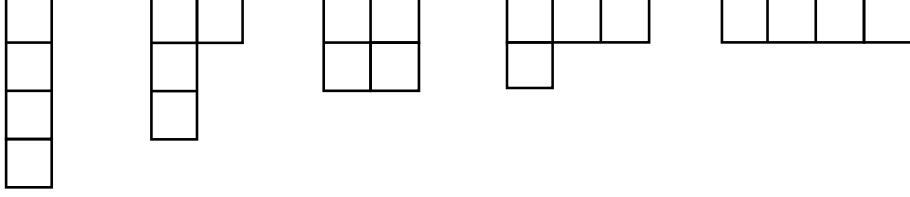
We can immediately notice that the symmetries of C require that all but the totally symmetric tensor S and one mixed symmetry tensor A must vanish. S has 15 independent components, and A has 6. Furthermore, A has corresponding symmetrizer $CS(x_1x_2)S(x_3x_4)A(x_1x_3)A(x_2x_4)$. Thus, A is exactly the tensor symmetric under permutation of i, j and k, l but antisymmetric under exchanges i, k and j, l.

Note that this mixed-symmetry subspace A corresponds to Backus' [**backus1970geometrical**] asymmetric tensor A. With S and A being defined component-wise as:

$$S_{ijkl} = \frac{1}{3}\big(C_{ijkl} + C_{ikjl} + C_{klij}\big)$$

$$A_{ijkl} = \frac{1}{3}\big(2C_{ijkl} - C_{ikjl} - C_{klij}\big)$$

The fully symmetric part can of course be converted to spherical harmonic components via a Clebsch-Gordon expansion. The mixed symmetry A, however requires further decomposition.

All six components can be described by the symmetric (but not traceless) tensor t, defined as the double partial contraction of A with the totally antisymmetric tensor ϵ as:

$$t_{ij} = \epsilon_i^{mk}\epsilon_j^{nl}A_{mnkl}$$

which can be reconstructed as the subtensor N as:

$$N_{ijkl} = \frac{1}{2}\big(\epsilon\epsilon - \epsilon\epsilon\big)t_{mn}$$

This tensor t then has a harmonic decomposition according to the order-two Clebsch-Gordon transformation between the J_z basis and the harmonic basis y_ℓ^m.

C $SO(3)$ Equivariant Layers

Since compositions of equivariant functions are themselves equivariant, we need only consider a set of basic equivariant functions which may then be composed

22

layer-wise to construct arbitrarily large equivariant networks. Below, we give brief descriptions of four common equivariant functions [**equivariant˙cohen**] used in such networks, namely: point-wise nonlinearities; self-interaction; pooling functions; and the heart of equivariant networks, equivariant convolution.

$SO(3)$-Equivariant Convolution

Features and filters in $SO(3)$ networks are both associated with $SO(3)$ representations, or spherical harmonics. Tensor products of these two representation spaces are equivariant under transformation of the two subspaces, i.e.:

$$\mathcal{D}^V \otimes \mathcal{D}^W = \mathcal{D}^{V \otimes W}.$$

By way of Clebsch-Gordan coefficients $c_{\ell_1 m_1 \ell_2 m_2}^{\ell_3 m_3}$, these tensor products of $SO(3)$ irreducible representations may be related to a third set of irreducible representations as [**sakurai**]:

$$(u \otimes v)_{\ell_o}^{m_o} = c_{\ell_1 m_1 \ell_2 m_2}^{\ell_o m_o} u_{\ell_1}^{m_1} v_{\ell_2}^{m_2}$$

where u and v are harmonic vectors of order ℓ_1 and ℓ_2, respectively. Since tensor products of representations are naturally equivariant, we may define $SO(3)$-equivariant convolution to be the scaled tensor product of the two representation spaces (i.e. that of the input feature space, and the filter space). Specifically, layer to layer convolutional maps $\mathcal{L}$ may be defined component-wise as[**tensorfieldnetworks**]:

$$\mathcal{L}_{acm_o}^{\ell_o}(\vec{r}_a, V_{acm_i}^{\ell_i}) = \sum_{m_f, m_i} c_{\ell_i m_i \ell_f m_f}^{\ell_o m_o} \sum_b F_{cm_f}^{\ell_f \ell_i}(r_{ab}) V_{bcm_i}^{\ell_i}$$

where the filter function $F_{cm_f}^{\ell_f \ell_i}(r_{ab})$ depends only on the distance between point a and b (as opposed to directional dependence, to maintain equivariance). Instances of such functions then generally have independent, trainable parameters for different rotational orders ℓ_f, ℓ_i, azimuthal orders m, and channels c.

Self-Interaction

Feature sets for individual objects may also update according to themselves as long as they act across m for every ℓ and only update according to the different channels c. That is, functions of the form:

$$V_{acm}^{\ell} \to \sum_c W_{c'c}^{\ell} V_{acm}^{\ell}$$

are also equivariant. Such functions we refer to as *self-interaction* layers since they act node-wise such that no information is exchanged between nodes.

Non-Linearities

We can also apply point-wise non linearities and maintain equivariance, as long as they only act on the $\ell = 0$ (that is, the rotationally-invariant) features of each layer. Non-linearities applied to features with index $\ell > 0$ break equivariance since rotations affect the values of specific components indexed by m.

Pooling

Pooling, or aggregation, across all elements or objects (index a) while preserving the m and ℓ indices is also equivariant. So functions of the form:

$$M_{cm}^{\ell} = \mathrm{AGG}_a(\{V_{acm}^{\ell}\})$$

where AGG is an arbitrary aggregation function performed only over the object index a, are also available in the construction of $SO(3)$ networks. Typically, pooling layers are utilized in the final layer of models to generate global representations of material structures from node-wise, atomistic, representations.

D   Real Spherical Harmonics

The (complex) spherical harmonics Y_{ℓ}^m can be transformed into a set of real spherical harmonics $Y_{\ell m}$ according the following relations:

$$Y_{\ell m} = \begin{cases} \frac{i}{\sqrt{2}}\left(Y_{\ell}^{-|m|} - (-1)^m Y_{\ell}^{|m|}\right): & m < 0 \\ Y_{\ell}^0: & m = 0 \\ \frac{1}{\sqrt{2}}\left(Y_{\ell}^{-|m|} + (-1)^m Y_{\ell}^{|m|}\right): & m > 0 \end{cases}$$

(of course, this choice is not entirely unique, with the choice made here assuming a Condon-Shortley phase included in the definition of Y_{ℓ}^m).

E   Formulas for Scalar Elastic Properties from Elastic Tensor

$$K_V = [(C_{11} + C_{22} + C_{33}) + 2(C_{12} + C_{23} + C_{31})]/9$$

$$K_R = 1/[(s_{11} + s_{22} + s_{33}) + 2(s_{12} + s_{23} + s_{31})]$$

$$K_H = (K_V + K_R)/2$$

$$G_V = [(C_{11} + C_{22} + C_{33}) - (C_{12} + C_{23} + C_{31}) + 3(C_{44} + C_{55} + C_{66})]/15$$

$$G_R = 15/[4(s_{11} + s_{22} + s_{33}) - 4(s_{12} + s_{23} + s_{31}) + 3(s_{44} + s_{55} + s_{66})]$$

$$G_H = (G_V + G_R)/2$$

$$E = 9K_H G_H/(3K_H + G_H)$$

G Distribution of Data

There were 3292 piezoelectric tensors, 7273 dielectric tensors, and 10286 elasticity tensors in the datasets used. This is after pruning all those available on Materials Project [**matproj**] to those with spherical harmonic components below some high-end cutoff, here 5000.

H Symmetry Classes of Tensors

H.1 Dielectric Tensors

The symmetry and number of independent components of dielectric tensors are determined by the crystal class of the associated structure.

$$\textbf{Triclinic}: \begin{bmatrix} a & d & e \\ & b & f \\ * & & c \end{bmatrix}$$

$$\textbf{Monoclinic}: \begin{bmatrix} a & 0 & d \\ & b & 0 \\ * & & c \end{bmatrix}$$

$$\textbf{Orthorhombic}: \begin{bmatrix} a & 0 & 0 \\ & b & 0 \\ * & & c \end{bmatrix}$$

$$\textbf{Tetragonal}, \textbf{Trigonal}, \textbf{Hexagonal}: \begin{bmatrix} a & 0 & 0 \\ & a & 0 \\ * & & b \end{bmatrix}$$

$$\textbf{Cubic}: \begin{bmatrix} a & 0 & 0 \\ & a & 0 \\ * & & a \end{bmatrix}$$

Component-wise MAE by Crystal Class
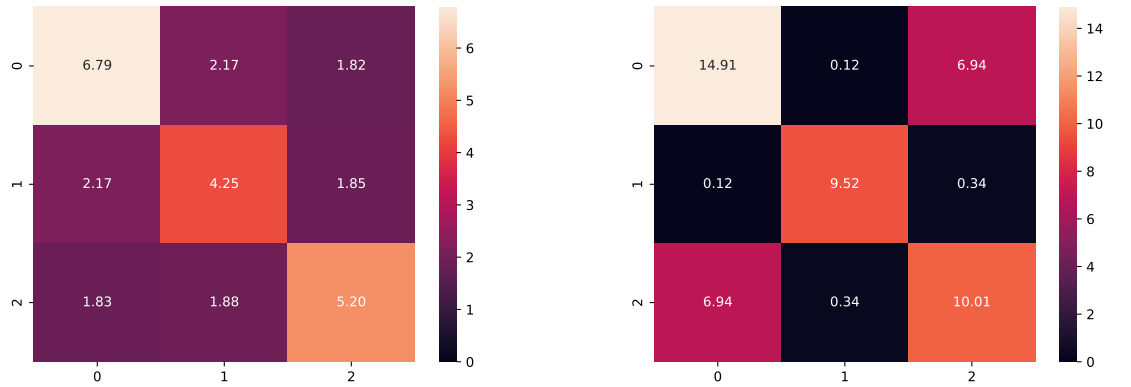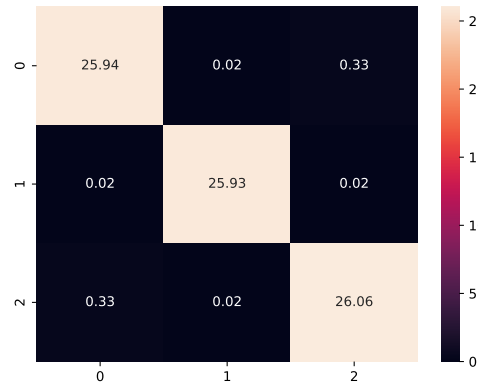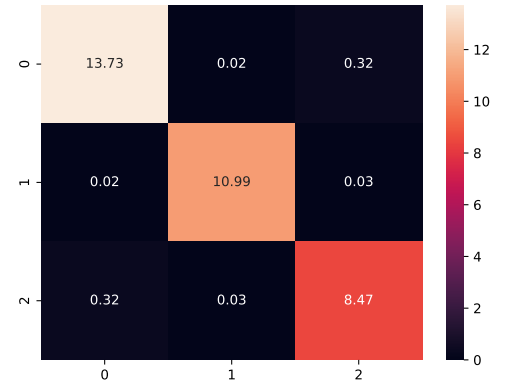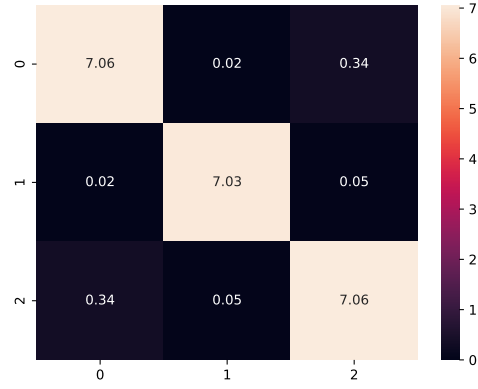
H.2 Piezoelectric Tensors

The symmetry and number of independent components of piezoelectric tensors are determined by the point group of the associated structure.

Component-wise MAE by Point Group
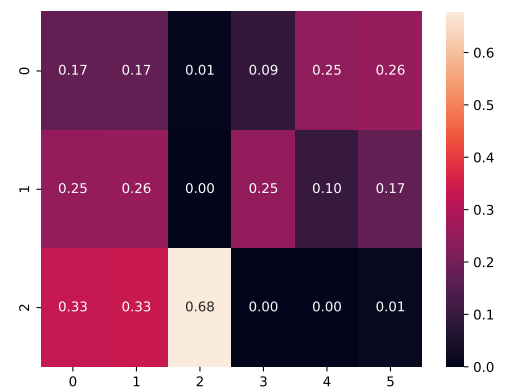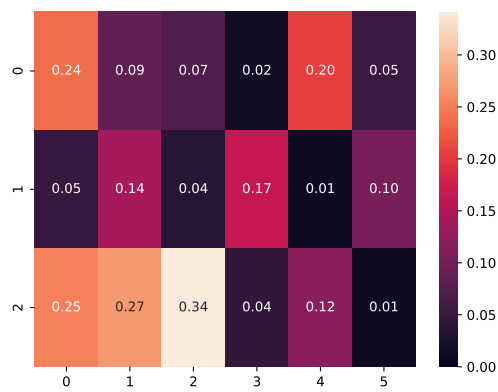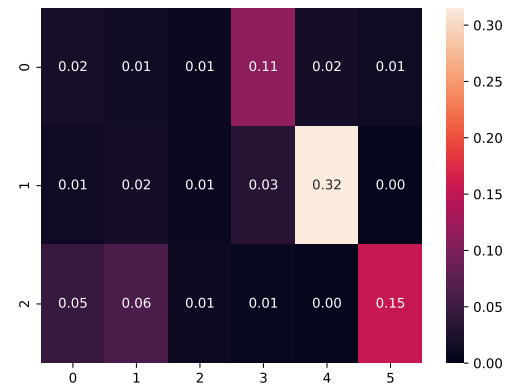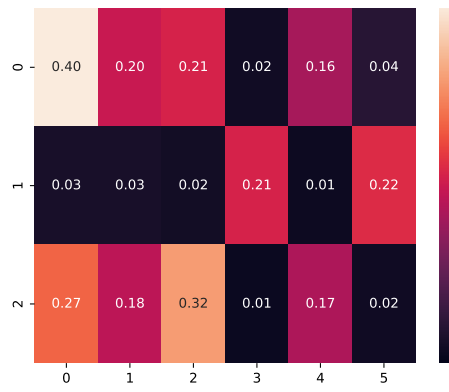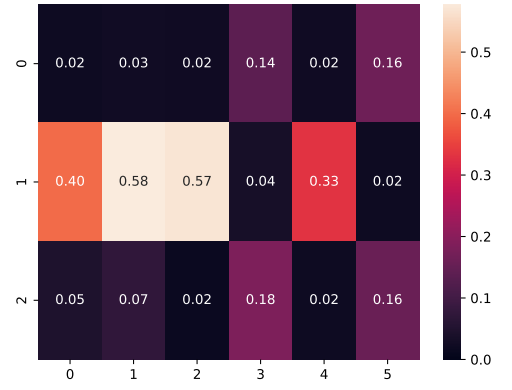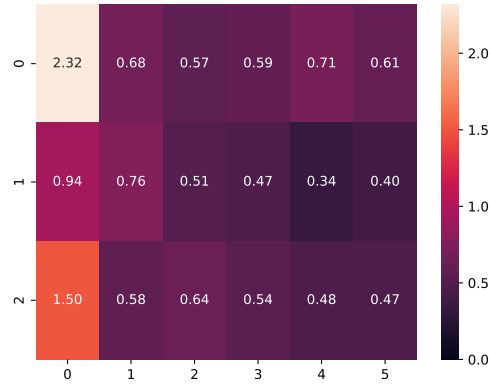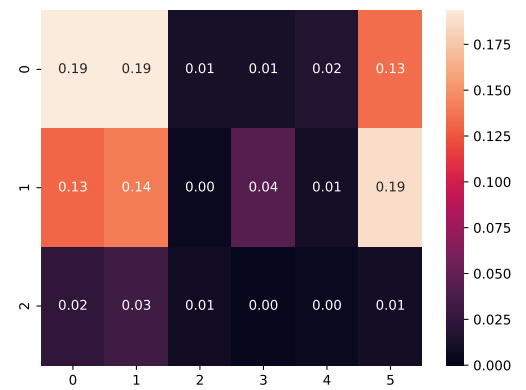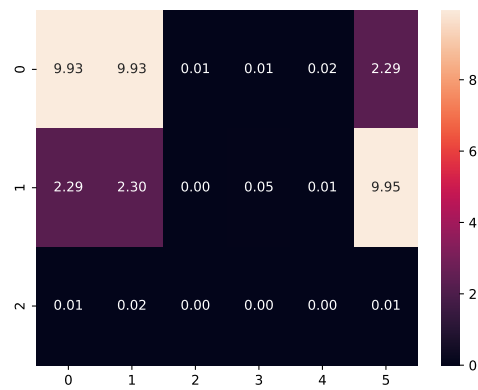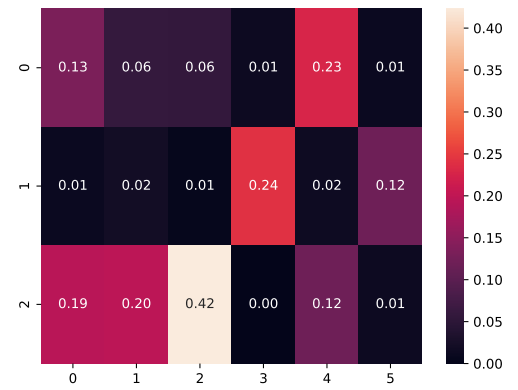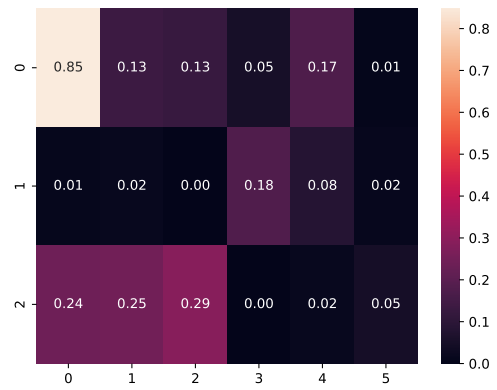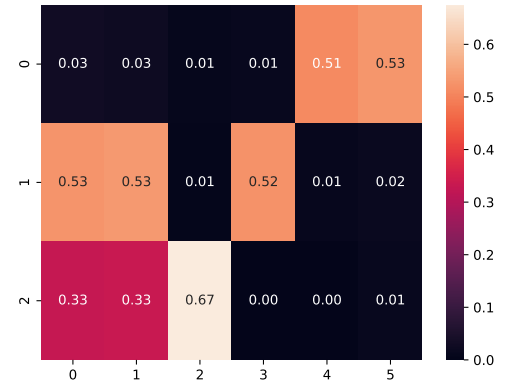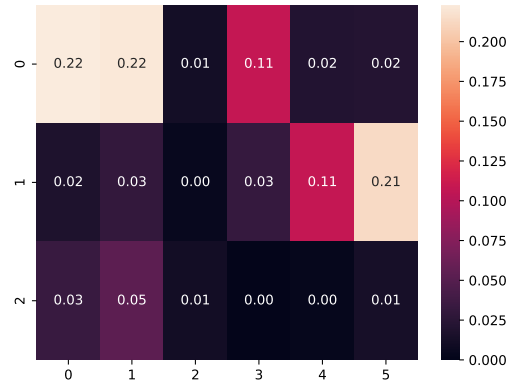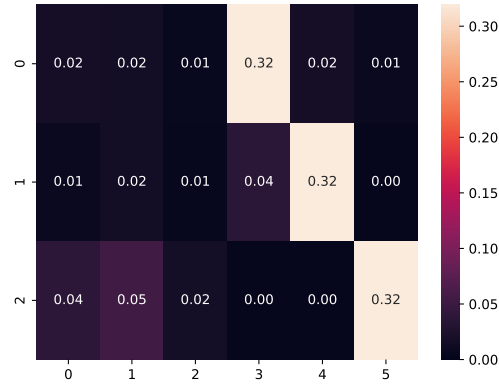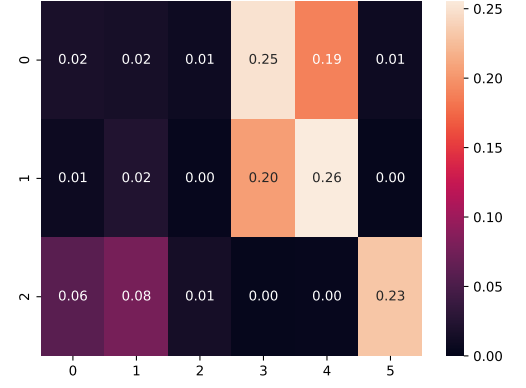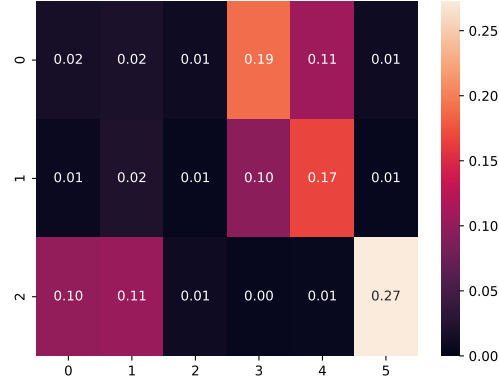
H.3   Elastic Tensors

The symmetry and number of independent components of elastic tensors are determined by the crystal class of the associated structure.

Component-wise MAE by Crystal Class
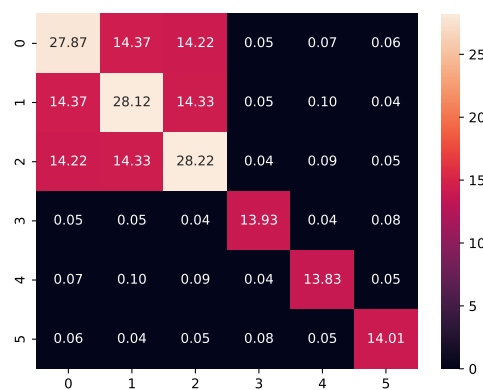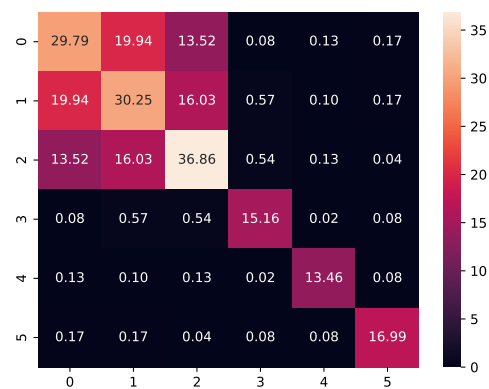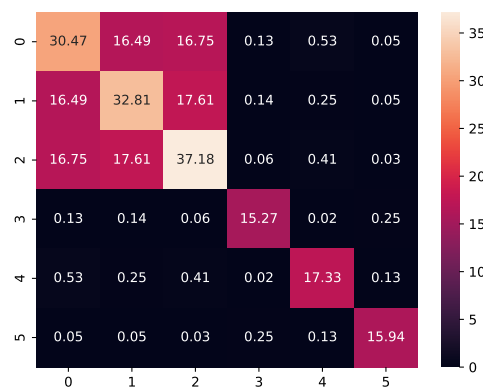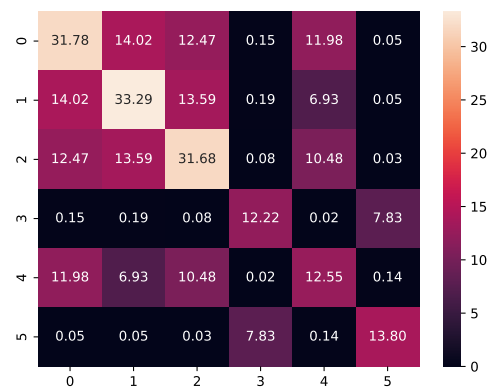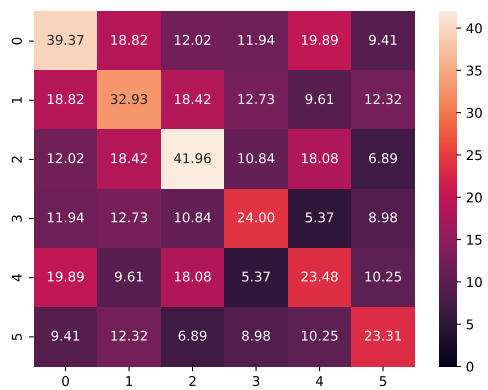
I   Elastic Tensor Component-wise MAE

I.1   Elastic Tensor Prediction from Band Gap
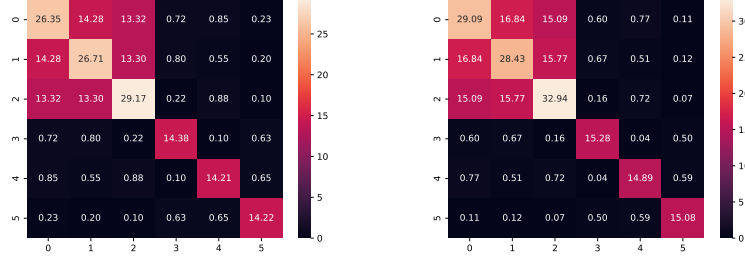
(a) SEConv Component-wise Performance
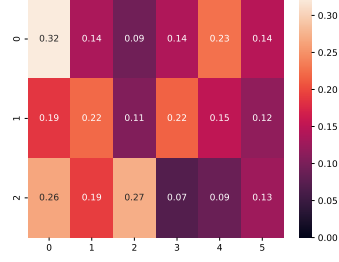

(b) SETransformer Component-wise Performance
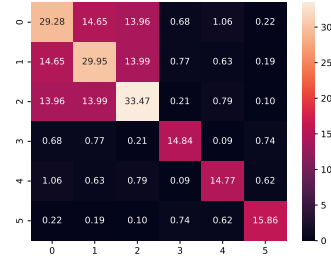
(c) SEGNN Component-wise Performance
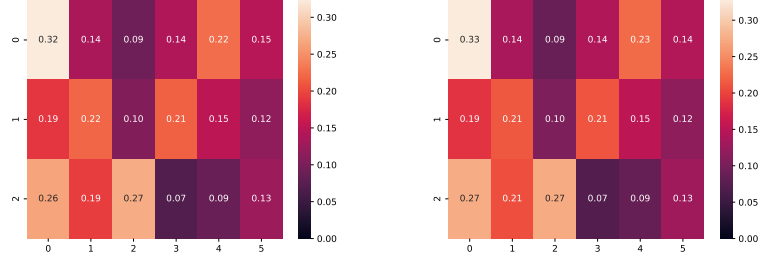
I.2 Elastic Tensor Prediction from Formation Energy
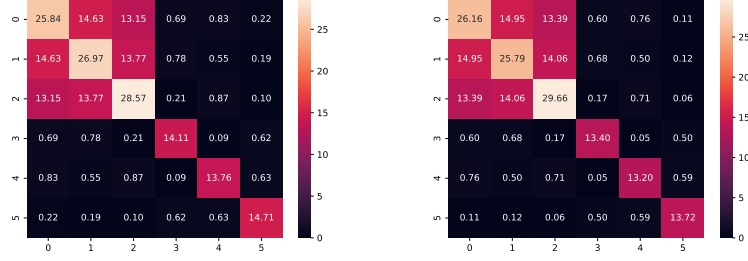
(a) SEConv Component-wise Performance
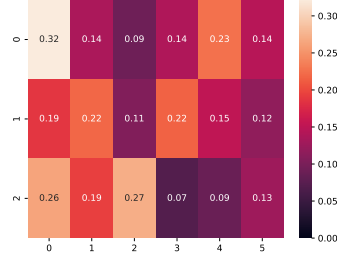

(b) SETransformer Component-wise Performance

(c) SEGNN Component-wise Performance

J   Dielectric Tensor Component-wise MAE

J.1   Dielectric Tensor Prediction from Band Gap
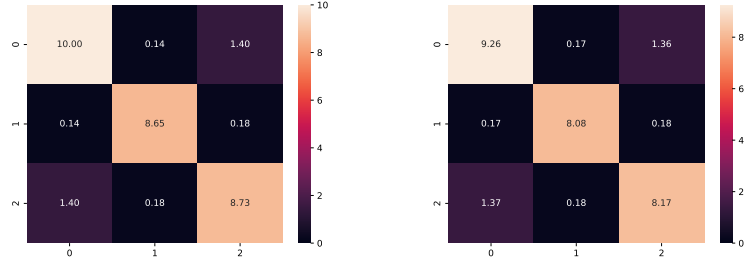
(a) SEConv Component-wise Performance

(b) SETransformer Component-wise Performance
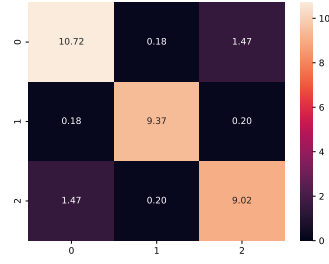
(c) SEGNN Component-wise Performance

J.2 Dielectric Tensor Prediction from Formation Energy
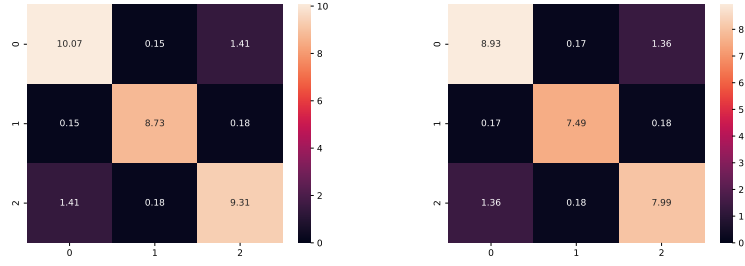
(a) SEConv Component-wise Performance

(b) SETransformer Component-wise Performance
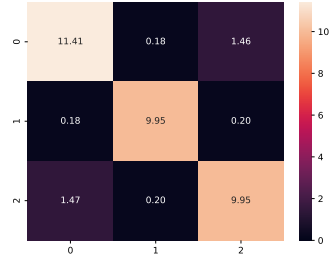
(c) SEGNN Component-wise Performance

J.3 Dielectric Tensor Prediction From Experiment 1

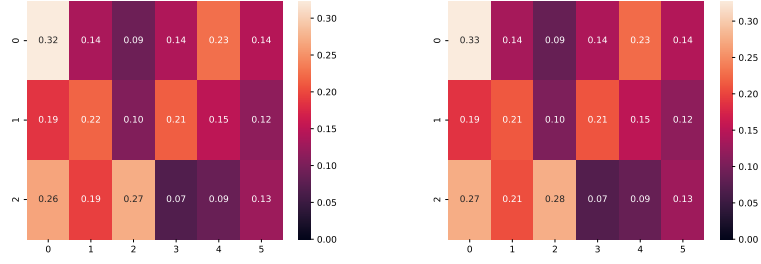### J.3.1 Band Gap as the Scalar Pretraining
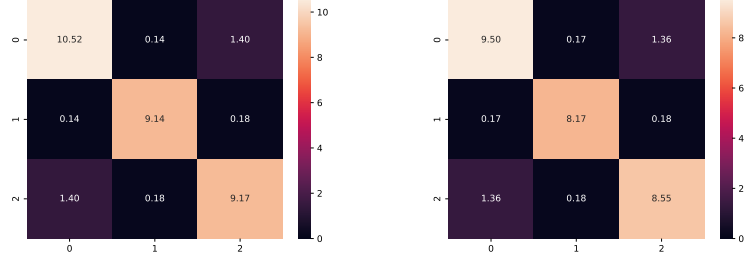
(a) SEConv Component-wise Performance

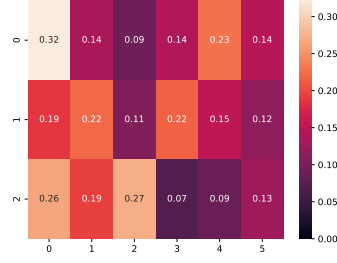(b) SETransformer Component-wise Performance
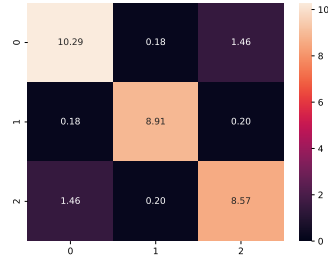
(c) SEGNN Component-wise Performance

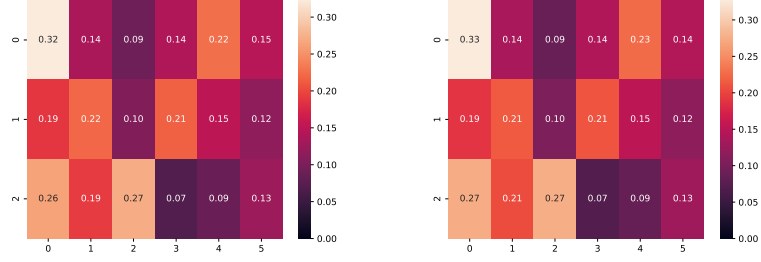### J.3.2 Formation Energy as the Scalar Pretraining
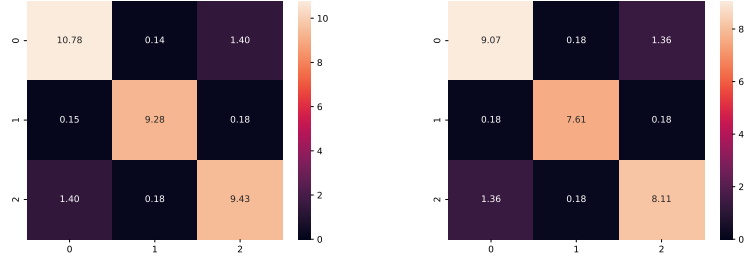
(a) SEConv Component-wise Performance
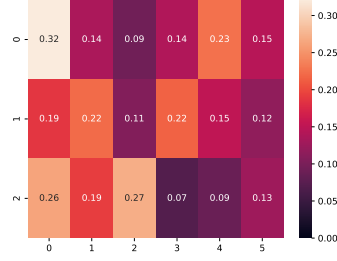

(b) SETransformer Component-wise Performance
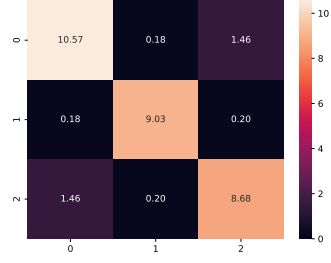
(c) SEGNN Component-wise Performance

K   Piezoelectric Tensor Component-wise MAE

K.1   Piezoelectric Tensor Prediction from Band Gap
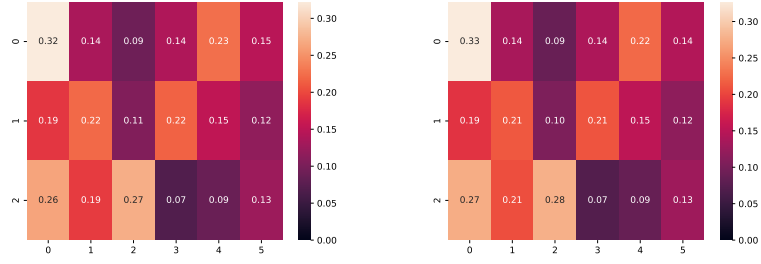
(a) SEConv Component-wise Performance

(b) SETransformer Component-wise Performance
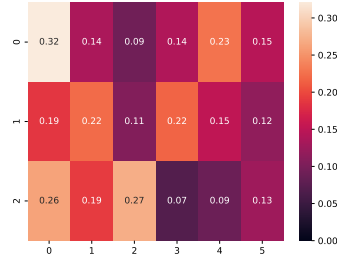
(c) SEGNN Component-wise Performance

K.2 Piezoelectric Tensor Prediction from Formation Energy
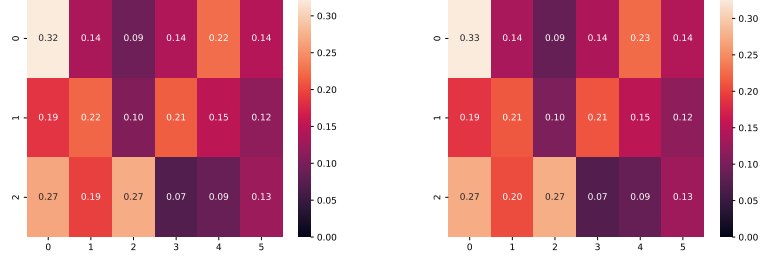
(a) SEConv Component-wise Performance

(b) SETransformer Component-wise Performance

(c) SEGNN Component-wise Performance

K.3 Piezoelectric Tensor Prediction From Experiment 1
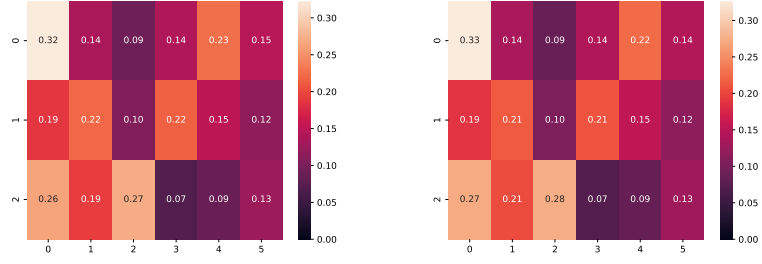
K.3.1 Band Gap as the Scalar Pretraining

(a) SEConv Component-wise Performance

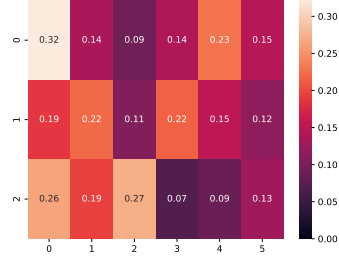(b) SETransformer Component-wise Performance

(c) SEGNN Component-wise Performance

K.3.2   Formation Energy as the Scalar Pretraining
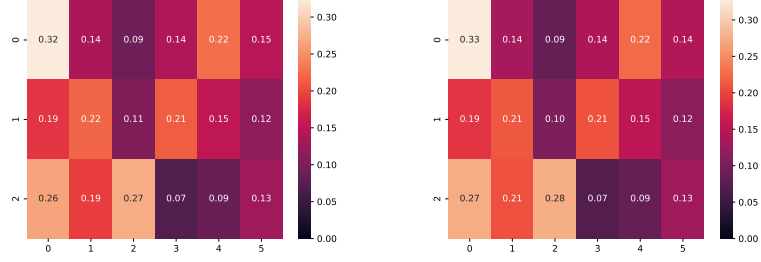

(a) SEConv Component-wise Performance
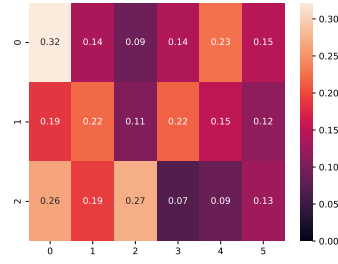

(b) SETransformer Component-wise Performance

(c) SEGNN Component-wise Performance

K.4 Piezoelectric Tensor Prediction From Experiment 2

K.4.1 Band Gap as the Scalar Pretraining

(a) SEConv Component-wise Performance

(b) SETransformer Component-wise Performance

(c) SEGNN Component-wise Performance

K.4.2 Formation Energy as the Scalar Pretraining

(a) SEConv Component-wise Performance

(b) SETransformer Component-wise Performance

(c) SEGNN Component-wise Performance

K.5   Piezoelectric Tensor Prediction From Experiment 3
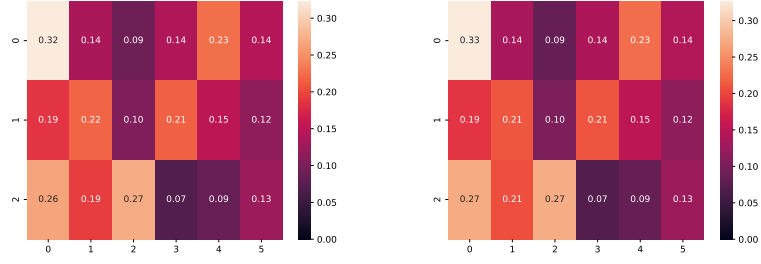
K.5.1   Band Gap as the Scalar Pretraining

(a) SEConv Component-wise Performance

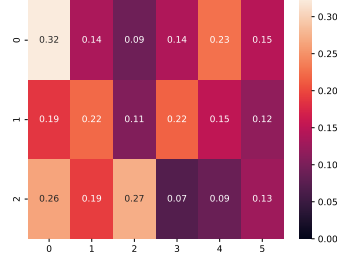(b) SETransformer Component-wise Performance

(c) SEGNN Component-wise Performance

K.5.2 Formation Energy as the Scalar Pretraining

(a) SEConv Component-wise Performance

(b) SETransformer Component-wise Performance

(c) SEGNN Component-wise Performance

F Table of Coefficients for $J_z \to y_\ell^m$ Transformation

order 2

	a_{00}	a_{+-}	$a_{0\pm}$	$a_{\pm\pm}$
y_0^0	1	-2		
y_2^0	1	1		
$y_2^{\pm1}$			$\sqrt{3}$	
$y_2^{\pm2}$				$\sqrt{\tfrac{3}{2}}$

order 3

	a_{000}	a_{0+-}	$a_{00\pm}$	$a_{+-\pm}$	$a_{0\pm\pm}$	$a_{\pm\pm\pm}$
y_1^0	3	-6				
y_3^0	$\tfrac{15}{7}$	$-\tfrac{5}{7}$				
$y_1^{\pm1}$			$\sqrt{\tfrac{3}{2}}$	$-2\sqrt{\tfrac{3}{2}}$		
$y_3^{\pm1}$			$2\sqrt{\tfrac{3}{2}}$	$\sqrt{\tfrac{3}{2}}$		
$y_3^{\pm2}$					$\sqrt{\tfrac{15}{2}}$	
$y_3^{\pm3}$						$\sqrt{\tfrac{5}{2}}$

Elastic (order 4)

(m = 0)

	a_{0000}	a_{00+-}	a_{0+0-}	a_{+-+-}	a_{++--}
y_0^0	1	2	-6	1	3
y_2^0	1	2	-3	1	-3
y_4^0	1	2	4	1	$\tfrac{1}{2}$

(m = 1, 2)

	$a_{000\pm}$	$a_{+-0\pm}$	$a_{\mp0\pm\pm}$	$a_{00\pm\pm}$	$a_{+-\pm\pm}$	$a_{\mp\pm\mp\pm}$
$y_2^{\pm1}$	$\sqrt{3}$	$\sqrt{3}$	$-3\sqrt{3}$			
$y_4^{\pm1}$	$2\sqrt{\tfrac{5}{2}}$	$2\sqrt{\tfrac{5}{2}}$	$\sqrt{\tfrac{5}{2}}$			
$y_2^{\pm2}$				$-2\sqrt{\tfrac{3}{2}}$	$-2\sqrt{\tfrac{3}{2}}$	$3\sqrt{\tfrac{3}{2}}$
$y_4^{\pm2}$				$\sqrt{\tfrac{5}{2}}$	$\sqrt{\tfrac{5}{2}}$	$2\sqrt{\tfrac{5}{2}}$

(m = 3, 4)

	$a_{\pm\pm\pm0}$	$a_{\pm\pm\pm\pm}$
$y_4^{\pm3}$	$\sqrt{\tfrac{35}{2}}$	
$y_4^{\pm4}$		$\tfrac{1}{2}\sqrt{\tfrac{35}{2}}$

Table 5: Coefficients for transformation between spherical basis components $a_{\alpha\beta\gamma\delta}$ and harmonic components y_l^m for a order-2,3,4 symmetric tensor. Note that the order-4 is only relevant for symmetric components of elastic tensors and correspond to a transformation [**mochizuki1988spherical**] between the J_z basis and their symmetric tensor s.

Elasticity Components (Mean ± STD in GPa)					
177.79 ± 134.19	72.27 ± 54.25	69.36 ± 53.86	0.02 ± 3.89	0.05 ± 4.12	0.02 ± 1.83
72.27 ± 54.25	178.16 ± 134.71	69.32 ± 53.86	-0.01 ± 4.18	0.08 ± 2.89	0.02 ± 2.02
69.36 ± 53.86	69.32 ± 53.86	175.33 ± 142.22	-0.02 ± 1.74	0.06 ± 3.91	-0.0 ± 1.06
0.02 ± 3.89	-0.01 ± 4.18	-0.02 ± 1.74	53.37 ± 47.03	0.0 ± 0.68	0.08 ± 2.88
0.05 ± 4.12	0.08 ± 2.89	0.06 ± 3.91	0.0 ± 0.68	53.86 ± 47.23	0.04 ± 3.82
0.02 ± 1.83	0.02 ± 2.02	-0.0 ± 1.06	0.08 ± 2.88	0.04 ± 3.82	55.74 ± 47.23

Piezoelectric Components (Mean ± STD in C/m^2)					
0.007 ± 1.382	-0.0149 ± 0.6281	-0.0207 ± 0.5903	0.0124 ± 0.4783	0.0186 ± 2.5957	-0.0163 ± 0.7636
-0.0084 ± 1.0081	0.0384 ± 1.0524	0.0088 ± 0.6786	0.0515 ± 2.2186	0.0078 ± 0.408	0.0048 ± 0.4732
-0.0099 ± 1.2187	-0.0162 ± 1.0639	-0.003 ± 1.0599	0.0002 ± 0.3991	-0.0084 ± 0.4473	0.0072 ± 0.7069

Dielectric Components (Mean ± STD)		
21.89 ± 72.55	0.02 ± 4.46	0.38 ± 18.24
0.02 ± 4.46	21.62 ± 71.08	-0.01 ± 4.04
0.38 ± 18.24	-0.01 ± 4.04	21.57 ± 85.67

Table 6: Component-wise mean and standard deviation for tensorial datasets used in this paper.