Symmetry Group Equivariant Neural Networks for Physical Systems

Alexander J. Heilman

May 6, 2025

Abstract

Contents

1 Literature Review & Background

Machine learning techniques have proven to be effective and efficient in many applications to the natural sciences. Particularly, supervised learning techniques applied to the prediction of material properties from (usually DFT-relaxed or idealized) atomic structures are of interest here.

1.1 Graph Neural Networks Applied to Atomic Systems

Graph neural networks act on input feature vectors associated with a corresponding graph, composed of nodes and edges between them. In applications to physical systems, we encode atomic arrangements (such as crystals and molecules) as graphs with associated features and then update their associated features, initially derived from physical properties, through a graph neural network.

Convolution on graphs (or "message passing" on graphs [?]) is a generalization of typical convolutional networks applied to image-related tasks, such as image classification. In typical convolutional networks, feature maps describing each pixel are updated by the features of their neighborhood of pixels (up to some set size). Through layers of convolution on graph features, features are also updated according to their neighbor's features, with neighbors here defined as nodes sharing edges.

In typical convolutional networks, pixels have fixed distances between them, allowing for discretized 'filters' that expect fixed distances between neighborhoods of pixels. However, in atomic systems, distances between atoms are continuous and thus require filters sensitive to continuously-distributed features describing these features. Continuous filters for such applications to atomic systems was first introduced in SchNet [?].

The widely-cited Crystal Graph Convolutional Neural Network (CGCNN [**?**]) then used continuous filters to great effect on a wide variety of material property predictions of crystalline systems.

1.2 $O(3)$-Equivariant Networks

Graph networks, in their application to atomic systems, are typically coordinate-system invariant by design. That is, the encoding of atomic information and arrangement typically is done in such a way without coordinate-system dependent quantities, instead relying on connectivity and inter-atomic distance to enforce invariance.

$O(3)$-Equivariant networks provide a framework in which direction quantities are encoded and, further, respected through convolution; in this way, outputs themselves may be treated as directional quantities. In general, 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 [**?**]. 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 [**?**]. These consist of: $SO(3)$-feature convolutions, ℓ-wise self-interactions and non-linearities, and pooling across nodes.

Through convolution, these filters and functions interact through tensor products, since 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 $c^{\ell_3 m_3}_{\ell_1 m_1 \ell_2 m_2}$, these tensor products of $SO(3)$ irreducible representations may be related to a third set of irreducible representations as [**?**]:

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

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[**?**]:

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

where the filter function $F^{\ell_f \ell_i}_{cm_f}(r_{ab})$ 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.

2 Previous Work

My previous work pertains to both (invariant) graph neural networks, as well as to equivariant networks in their application to materials science. Particularly, these works focus on improving the expressiveness of invariant graph neural networks by generalizing descriptors to hypergraphs; and utilize $O(3)$-equivariant networks to novelly predict tensorial quantities. Below, the techniques of both of these works are detailed.

2.1 Crystal Hypergraph Networks

The primary problem with the representations of material systems in invariant graph neural networks is that they often lack 'geometrical resolution' beyond interatomic distance. For example, in the proto-typical techniques of CGCNN, distinct coordination environments of the same interatomic-distance and coordination number will be mapped onto equivalent graphs.

Several approaches have been taken to increase the geometric resolution of invariant graph neural networks: namely, the inclusion of bond-angles as additional triplet features, dihedral angles as quartic features, etc.; and, alternatively, the explicit inclusion of higher-order (than strictly pair-wise) tailored neighborhood features describing common motifs in crystal systems, or more generally, common coordination environments.

My approach to increasing geometric resolution, while maintaining invariance, was to extend the graphical descriptions of material systems to hypergraph descriptions, in which edges (now hyperedges) are then allowed to contain more than two atoms. This allows for all invariant features to be naturally associated with arbitrary order structures in materials, simultaneously on equal footing. For example, pairs, triplets, and coordination environments may all be described in the same crystal hypergraph as hyperedges of differing orders.

Of course, these different structures require different sets of descriptors, and thus inequivalent feature vectors, suggesting these crystal hypergraphs are naturally heterogeneous in their hyperedges. That is, not all hyperedges in crystal hypergraphs are of the same type. This simply means we need different convolutional filters for each type of hyperedge.

2.2 Equivariant Tensor Predictions

Spherical harmonics, the basis vectors of $O(3)$-invariant subspaces, naturally correspond to sets of basis tensors as well. This correspondence allows for the direct prediction of tensorial properties from the output of $O(3)$ networks in a way that respects transformations on the input coordinate system, acting essentially as similar transformations on the output tensors. In this way, $SO(3)$ 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. In this work, we predicted three common material properties described by tensors, namely: the dielectric tensor for anisotropic materials; the piezoelectric tensor, relating strain to internal electric fields; and the elasticity tensor, relating strain to stress.

The first step in such applications is the decomposition of tensor spaces into ℓ and m indexed subspaces (that is, into irreducible $SO(3)$ tensor subspaces). 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.

3 Proposed Project

$SE(3)$-Equivariant networks have proven effective in applications within the physical sciences, suggesting that physical symmetries are a useful inductive bias in the design of machine learning techniques. While $SE(3)$ networks, in principle, can fully incorporate all possible symmetries of a system, in some applications

other symmetry subgroups may be more relevant. One case of this is applications of machine learning to condensed matter systems, where atoms are arranged in periodic structures where each atomic position has some point symmetry group.

In $SE(3)$ Equivariant networks, initial edge features $\vec{e}_{ij}^{L=0}$ are often spherical harmonic functions $Y_m^\ell(\vec{r}_{ab}$ evaluated for interatomic radii $\vec{r}_{ab}$ of atom pairs (a, b). Point group features may be derived from interatomic radii by way of projection operators $\hat{P}_{ij}^\mu$ which project arbitrary vector spaces onto subspaces V_μ transforming as irreducible representations μ of point group G, and defined as:

$$\hat{P}_\alpha^{kk} = \frac{d_\alpha}{N} \sum_g \left[\Gamma_{kk}^\alpha(g)\right]^* O(g)$$

where d_α is the dimension of irrep Γ^α, N is the order of the group, and $O(g)$ is the operator representation acting on the feature space of relevance. While $SE(3)$ networks rely on Clebsch-Gordan expansions of tensor products of neighboring sites and filters, point group networks instead require generalized CG coefficients or 'coupling coefficients' $U_{\alpha i \beta j}^{\gamma n}$ for the direct sum decompositions of point group-irreducible features:

$$\psi_n^\gamma = \bigoplus_{\alpha i, \beta j} U_{\alpha i \beta j}^{\gamma n}\left(u_i^\alpha \otimes v_j^\beta\right)$$

With these coefficients, a similar convolutional structure may be adopted, as in $SE(3)$ networks, though now instead of harmonic indices ℓ, m for features, we have point group irreducible representation indices γ, n:

$$\left(v_{nc}^{L+1}\right)_n^\gamma = \left(v_{nc}^L\right)_n^\gamma + \sum_{b \in \mathcal{N}(n)} \sum_{\alpha i, \beta j} U_{\alpha i \beta j}^{\gamma n}\left(F_c^L(r_{nb})\right)_j^\beta \left(v_{bc}^L\right)_i^\alpha$$

We propose to incorporate the above techniques into a novel framework for point-group-equivariant graph networks, suited for applications where systems display particular symmetries reduced from the full rotation symmetry of $SO(3)$. One particularly relevant application is the prediction of condensed matter Hamiltonian elements. In solid-state systems based upon an underlying periodic lattice, each atomic position displays a certain point group symmetry. Thus the Hamiltonian of the system must share a simultaneous set of eigenvalues with the symmetry operations since in this case they necessarily commute. Essentially then, the Hamiltonians solutions must be indexed by the same set as the irreducible subspaces V_μ, each of dimension d_μ. The basis set of one-particle wavefunctions then must have a separable form:

$$\psi_{nm}^\mu(\vec{r}) = R_{\mu nm}(r)\Omega_n^\mu(\theta, \phi)$$

where $1 \leq n \leq d_\mu$. The radial function $R_{\mu nm}(r)$ then depends on the specific form of the Hamiltionian and may have a very complicated structure. The prediction of this functional dependence is thus a well suited for the machine learning applications discussed here. Such models provide a method by which we may predict Hamiltionian elements naturally associated with these point-group indexes, allowing for complete descriptions of interactions between neighboring sites irreducible subspaces without truncationss (such as ℓ_{max} in $SE(3)$ networks).

Furthermore, these models allow for the prediction of properties that respect the known symmetries of physical systems, by reading out features that transform as the trivial representation, allowing for more intuitive predictions of scalar (totally invariant) quantities, as well as tensorial properties with desired physical symmetries.

The implementation of these models requires the availability of the analagous Clebsch-Gordan coefficients, $U_{\alpha i \beta j}^{\gamma n}$, which we refer to more generally as *coupling coefficients*, for all point groups of interest. These may be defined, from particular forms of representations $\Gamma^\alpha, \Gamma^\beta, \Gamma^\gamma$ for all point groups of interest from Dirl's formula [?]:

$$\left(U_{\alpha i \beta j}^{\gamma n}\right)^m = \sqrt{\frac{d_\gamma}{N_G}}\left(\sum_{g \in G} \Gamma_{qq}^\alpha(g)\Gamma_{ss}^\beta(g)\Gamma_{aa}^{\gamma\dagger}(g)\right)^{-\frac{1}{2}} \cdot \sum_{g \in G} \Gamma_{iq}^\alpha(g)\Gamma_{js}^\beta(g)\Gamma_{na}^{\gamma\dagger}(g)$$

It should be noted here that the definition of these coefficients thus requires explicit (fixed) forms for each irrep of each point group of interest.

4

4 Methods & Timeline