Lorentz Group Equivariant Jet Autoencoder

Zichun Hao

Department of Physics, University of California, San Diego

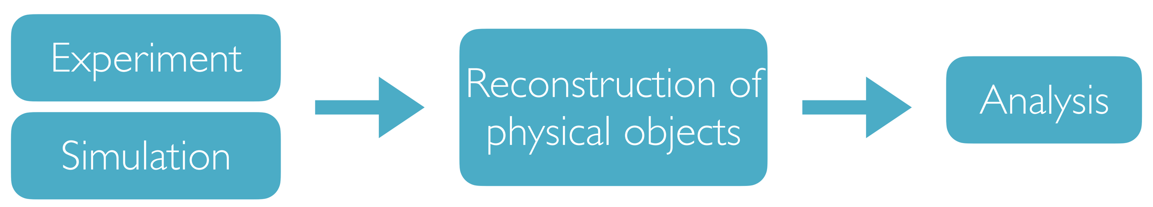
UC San Diego

Introductions

The goal of this project is to design a jet generative deep neural network model that is Lorentz group equivariant.

Particle Physics

Particle physics is the study of elementary particles, including their existences, intrinsic properties, and interactions. A very useful tool in experimental particle physics is **colliders**, in which beams of highly relativistic particles (i.e. with speed close to the speed of light) are collided together. **Jets**, collimated sprays of particles, can provide key insights into the nature of the underlying dynamics and interactions and therefore are the central object in the analyses of particle physics experiments. **Simulations** of collision events are an important part of data analysis in particle physics. Simulations translate the theoretical model into experimental signatures that can be used to construct physical objects from raw data. Classical simulation programs, such as `GEANT4`, are accurate but are slow, producing data at a rate $\mathcal{O}(1\ \mathrm{min/event})$.



Groups

A **group** is an ordered pair $(G, \star)$ where G is a set and $\star$ is a binary operation on G satisfying the following axioms:

- Closure: $\forall a, b \in G\ (a \star b \in G)$
- Associativity: $\forall a, b, c \in G\ ((a \star b) \star c = a \star (b \star c))$
- Existence of Identity element: $\exists e \in G\ \forall a \in G\ (a \star e = e \star a = a)$
- Existence of Inverse element: $\forall a \in G\ \exists a^{-1} \in G\ ((a \star a^{-1}) = a^{-1} \star a = e)$

We often just say that G is a group when the operation $\star$ is clear.

Groups Representations

A **representation** of a group G on a vector space V is a map $\rho : G \to \mathrm{GL}(V)$ such that $\rho(g_1 g_2) = \rho(g_1)\rho(g_2)$ for all $g_1, g_2 \in G$, where $\mathrm{GL}(V)$ is the set of $\dim V \times \dim V$ invertible matrices. Here, V is called the **representation space**. A representation is **reducible** if there exists a non-trivial invariant vector subspace $W \subsetneq V$ (and $W \neq \{\mathbf{0}\}$) with respect to the group element, that is $g\omega \in W$ for all $g \in G$ and $\omega \in W$. **Irreducible** representations (**irreps**) are not reducible.

Lorentz Group

The **Lorentz group**, SO(3,1), is a fundamental group in particle physics. It comprises of all possible compositions of spatial **rotations** and Lorentz **boosts** (the transformation of different inertial frames without a rotation which mixes space and time). Physical objects in a collision experiment should all live in the representation space of SO(3,1). Irreps of SO(3,1) can be labeled by (j^+, j^-), where $j^+, j^- = 0,\ 1/2,\ 1,\ 3/2,\ 2,\ \ldots$.

Equivariant Maps

$$\begin{tikzcd} V \arrow[r, "\rho(g)"] \arrow[d, "F"'] & V \arrow[d, "F"] \\ V' \arrow[r, "\rho'(g)"'] & V' \end{tikzcd}$$

Given two representations (V, ρ) and (V', ρ') of group G, a map $F : V \to V'$ is called **equivariant** if $F(\rho(g)) \cdot v) = \rho'(g) \cdot F(v)$ for $v \in V$ and $g \in G$. For a **Lorentz group equivariant map** $\rho : \mathbb{R}^{3,1} \to \mathbb{R}^{3,1}$, it should be true that $\rho(gp^\mu) = g\rho(p^\mu)$ for all $g \in \mathrm{SO}(3,1)$ and $p^\mu \in \mathbb{R}^{3,1}$.

Neural Networks

Deep Neural Network (DNN)

A **deep neural network** (DNN) is a machine learning algorithm that employs multiple hidden neural layers with structure. A DNN can be thought of as a map that takes an input and produces some output. In our case, both input and output are the 4-momenta of particles in a jet. DNN has shown its potential in dealing with a large amount of data, so it could be a good solution to slow generating rate of classical event simulation programs.

Autoencoder

This generative model is an **autoencoder**, which consists of an encoder and decoder. The **decoder** takes the real data as input and compresses its dimension down to the **latent space**. Then, the encoder takes the data in the latent space and attempts to reconstruct the data as much as possible. In this way, the model is forced to extract the most crucial features in the data. The decoder can then be used as a generate data. If the property of the latent space is well studied, one can make the model generate the desired data.

Architecture Overview

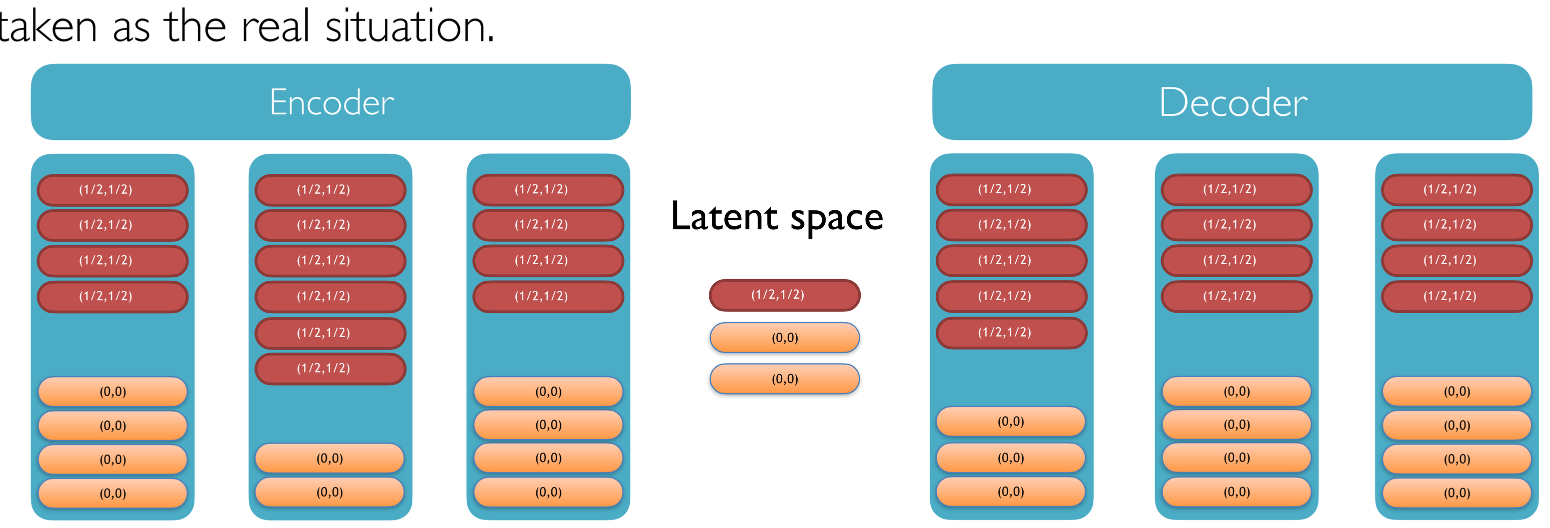
How to Achieve Group Equivariance?

The model works on the irreps of SO(1,3) to achieve group equivariance. The architecture is adapted from the **Lorentz Group Network** architecture built by Bogatskiy et al. Specifically, scalars live in the representation space of $(0,0) = D^0 \otimes D^0$ irrep, and the 4-momenta p^μ live in the representation space of the $(1/2,1/2) = D^{1/2} \otimes D^{1/2}$ irreps. The central part is the **CG layers**, each of which

- plays the role of an MLP in a normal autoencoder,
- is a **graph neural network** (neural network that has a graph structure),
  - with irreps as node features and learnable $f(p^2)$ as edge features (where p^2 is a scalar computed from the inner product of irreps),
- takes the tensor product of two irreducible representations and decompose to direct sums of irreps again using the **Clebsch-Gordan (CG) decomposition**,
  - where tensor product between two nodes models mutual interactions, and tensor product of a node with itself models the self-interactions,
- mixes the scalar using an extra MLP for better adaptability of the model, and
- mixes the updated irreps to wanted multiplicities using a learnable matrix,
  - In which only irreps with the same weights (j^+, j^-) are mixed.

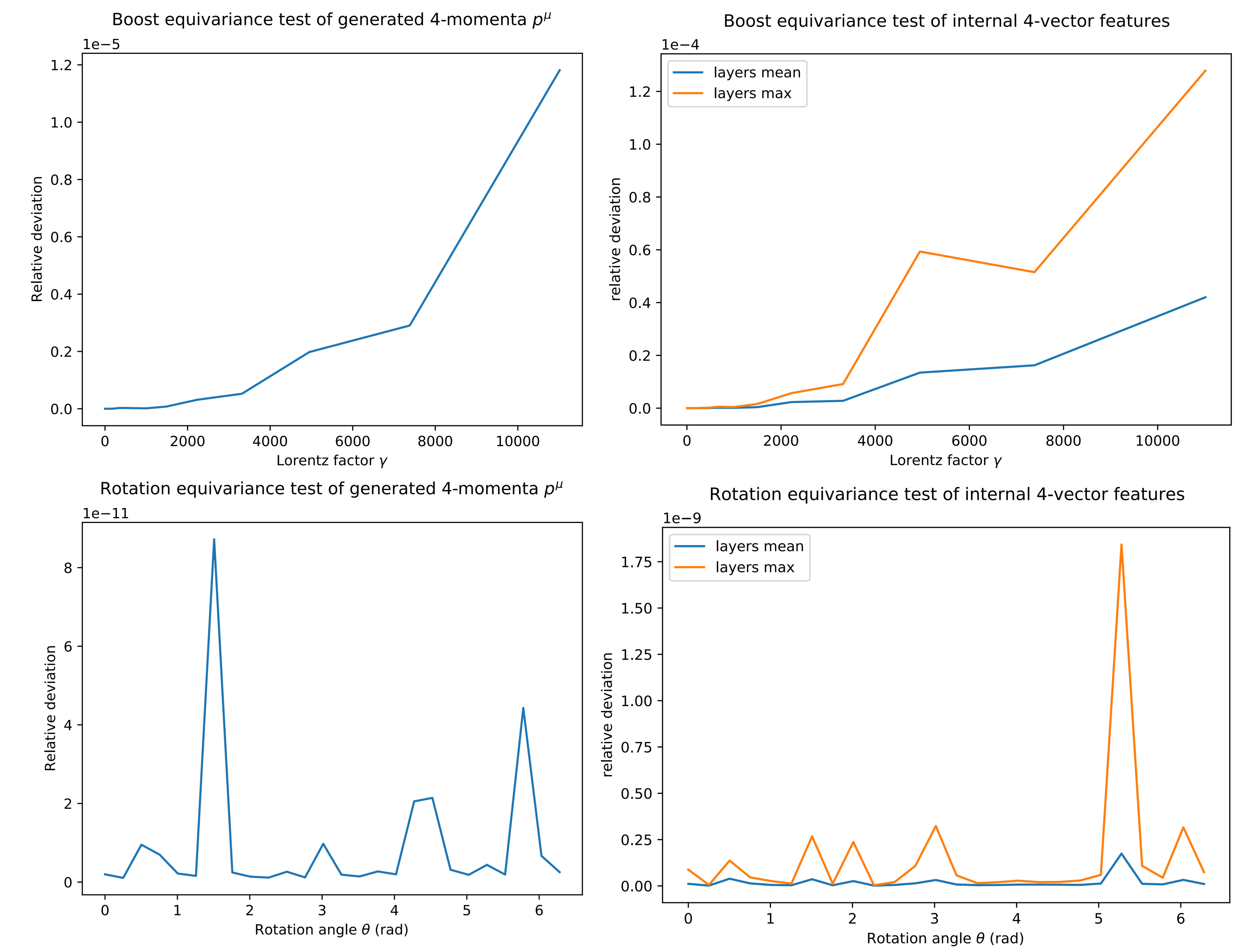
Schematic

In the diagram, $(1/2,1/2)$ represents 4-vectors, and $(0,0)$ represents scalars. The graph structure is ignored for simplicity. This diagram just shows the general idea of the architecture; the number of layers and the multiplicity of irreps in each layer should not be taken as the real situation.



Model Equivariance Tests

Boost and rotation equivariance tests are performed on the model using both outputs and the internal features. The input is boosted up to $\gamma = 11013.2$ and rotated up to $\theta = 2\pi$. Denote the neural network as $\mathrm{LGNAutoencoder}$, the transformation as $\Lambda \in \mathrm{SO}(3,1)$, and the jet data and internal 4-vector features as p. Then, $\mathrm{LGNAutoencoder}(\Lambda p)$ and $\Lambda \cdot \mathrm{LGNAutoencoder}(p)$ are compared against each other, and the element-wise relative deviation is computed against the expected output $\Lambda \cdot \mathrm{LGNAutoencoder}(p)$. The results are plotted using `numpy` and `mathplotlib`.



Acknowledgement

Special Thanks

I would like to express my sincere gratitude to my mentors Professor **Javier Duarte** and **Raghav Kansal**, and labmate **Farouk Mokhtar**, for giving tremendous helps and guidance to me and this project. Without their helps, making progress in this project would have been impossible. I would also like to thank the Faculty Mentor Program (**FMP**) for giving me the opportunity to conduct research and Dr. **Claire Kim** for keeping the program organized and patiently answer questions about the program.

References

- A. Bogatskiy et al., "Lorentz group equivariant neural network for particle physics", arXiv:2006.04780.
- D. Dummit and R. Foote, "Abstract Algebra". Wiley, 3rd edition, 2003.
- R. Sexl and K. Urbantke, "Relativity, Groups, Particles", Springer, 2001, doi: 10.1007/978-3-7091-6234-7.