Contribution ID: **34**   Type: **Regular talk**

Lorentz Equivariant Neural Networks for Particle Physics

Thursday 22 October 2020 14:00 (20 minutes)

We present a new set of neural network architectures, Lorentz group covariant architectures for learning the kinematics and properties of complex systems of particles. The novel design of this network, called LGN (Lorentz Group Network), implements activations as vectors that transform according to arbitrary finite-dimensional representations of the underlying symmetry group that governs particle physics, the Lorentz group. The nonlinearity of the network is based on the tensor product of representations of the Lorentz group. Consequently, the architecture is inherently covariant under Lorentz transformations and is capable of learning not only fully Lorentz-invariant objectives such as classification probabilities, but also Lorentz-covariant vector-valued objectives such as 4-momenta, while exactly respecting the action of the group. Imposing the symmetry leads to a significantly smaller ansatz (fewer learnable parameters than competing non-covariant networks), and potentially a much more interpretable model. To demonstrate the capability and performance of this network, we study the ability to classify systems of charged and neutral particles at the Large Hadron Collider resulting from the production and decay of highly energetic quarks and gluons. Specifically, we choose the benchmark task of classifying and discriminating jets formed from the hadronic decays of Lorentz-boosted massive particles from the background of light quarks and gluon jets. We show that we are able to achieve similar performance compared to other state-of-the-art neural networks trained to perform this classification task while also maintaining significantly broader generality regarding the structural origin of the physical processes involved. Moreover, we present simplified invariant and covariant architectures tailored to specific tasks with 4-vector inputs, which can be trained faster and more efficiently than the general LGN architecture.

Primary authors: BOGATSKIY, Alexander (University of Chicago); OFFERMANN, Jan Tuzlic (University of Chicago (US)); MILLER, David (University of Chicago (US)); ROUSSI, Marwah (University of Chicago); Prof. KONDOR, Risi (Flatiron Institute, University of Chicago)

Presenter: BOGATSKIY, Alexander (University of Chicago)

Session Classification: Workshop

Track Classification: 2 ML for analysis : Application of Machine Learning to analysis, event classification and fundamental parameters inference