Lorentz-Equivariant Geometric Algebra Transformers for High-Energy Physics

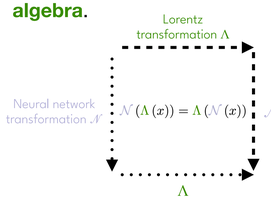
Jonas Spinner*
Víctor Bresó Pla*
Pim de Haan
Tilman Plehn
Jesse Thaler
Johann Brehmer

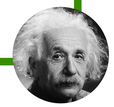
UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

Paper Code

Lorentz Equivariance

- **Equivariant networks** use knowledge about **symmetry** in their architectures, instead of having to learn it from data.
- The Lorentz or spacetime symmetry of special relativity is very important for **fundamental particle interactions**.
- It is easy to implement in a **geometric algebra**.

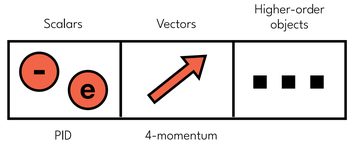




+

Geometric Algebra

- LHC data is structured as **particles**, specified by type information and a 4-momentum.
- The **geometric algebra** $\mathbb{G}_{1,3}$ is a natural representation for particles.
- Algebra elements are organized in **grades** with different geometric types including **higher-order objects**.



+

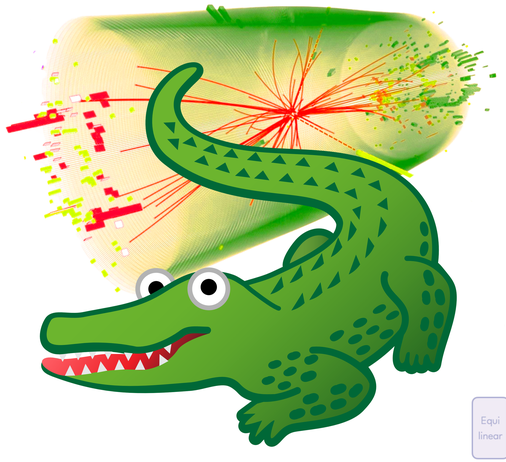
Transformer

- Particles are **permutation equivariant**.
- Transformers process point clouds with excellent **speed and memory scaling**.
- Transformer layers can become **equivariant** if the inputs, hidden states and outputs of the network are embedded in the **geometric algebra**.

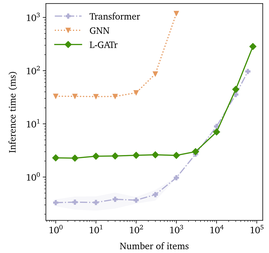
Layer type	Transformer	L-GATr
Linear(x)	$vx + w$	$\sum_{k=0}^{4} v_k (x)_k + \sum_{k=0}^{4} w_k \gamma^k (x)_k$
Attn.$(q, k, v)_{ic}$	$\sum_{j=1}^{n_c} \mathrm{Smax}_j \left(\sum_{c'=1}^{n_c} \frac{q_{ic'} k_{jc'}}{\sqrt{n_c}} \right) v_{jc}$	$\sum_{j=1}^{n_c} \mathrm{Smax}_j \left(\sum_{c'=1}^{n_c} \frac{\langle q_{ic'}, k_{jc'} \rangle}{\sqrt{16 n_c}} \right) v_{jc}$
LayerNorm(x)	$x \left[\frac{1}{n_c} \sum_{c=1}^{n_c} x_c^2 \right]^{-\frac{1}{2}}$	$x \left[\frac{1}{n_c} \sum_{c=1}^{n_c} \sum_{k=0}^{4} \left\lvert \langle (x_c)_k, (x_c)_k \rangle \right\rvert \right]^{-\frac{1}{2}}$
Activation(x)	GELU(x)	GELU$(\langle x \rangle_0) x$
GP(x, y)	–	xy

=

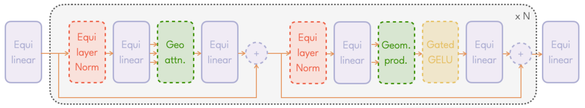
L GA Tr



- **L-GATr is a versatile and effective architecture for multiple tasks in the field of particle physics.**
- It supports partial and approximate symmetries through **symmetry-breaking inputs**.
- It is **faster and more memory efficient** than other equivariant baselines thanks to its transformer structure.



- It features a **standard transformer architecture** based on layers that process geometric algebra objects:



Background

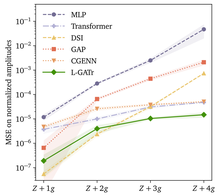
- At the **LHC**, fundamental properties of particles are studied by colliding highly accelerated protons.
- The LHC collects **petabytes of data every second**. Scientific insight is extracted from them by leveraging a complex **workflow** to compare simulations and measurements.

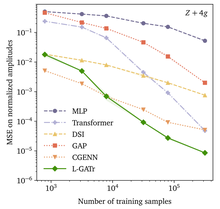
- The main representation of data at the LHC is the **particle**, characterized by scalar labels and a vector containing its energy and spatial momentum called the 4-momentum.
- The **Lorentz group** specifies the spacetime symmetry from Einstein's theory of special relativity, and it dictates how particles transform under rotations and boosts.

Experiments

Regression: QFT Amplitudes

- Exact but expensive task at the start of the LHC workflow. Neural surrogates are fast, but modeling high-multiplicity processes is challenging.
- We examine a process with an increasing number of particles and find **good scaling to high particle multiplicity.**





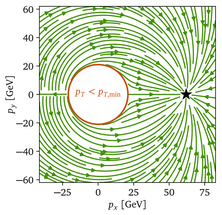
Classification: Top Tagging

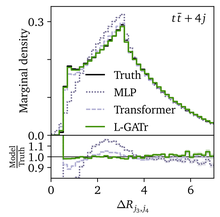
- The objective is to identify jets originated from the decay of a top quark. A reliable top tagger is critical to explore new physics scenarios.
- Collider topology breaks the Lorentz symmetry in this task. We introduce **reference vectors** to account for this.
- We observe that L-GATr is **on par with or better than the strongest baselines** on this well-studied benchmark.

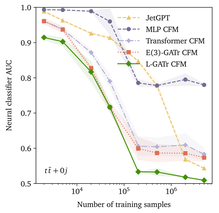
Model	Accuracy	AUC	$1/\epsilon_B$ ($\epsilon_S = 0.5$)	$1/\epsilon_B$ ($\epsilon_S = 0.3$)
TopoDNN	0.916	0.972	–	295 ± 5
LoLa	0.929	0.980	–	722 ± 17
P-CNN	0.930	0.9803	201 ± 4	759 ± 24
N-subjettiness	0.929	0.981	–	867 ± 15
PFN	0.932	0.9819	247 ± 3	888 ± 17
TreeNiN	0.933	0.982	–	1025 ± 11
ParticleNet	0.940	0.9858	397 ± 7	1615 ± 93
ParT	0.940	0.9858	413 ± 16	1602 ± 81
LorentzNet*	0.942	0.9868	498 ± 18	2195 ± 173
CGENN*	0.942	0.9869	500	2172
PELICAN*	**0.9426** ± 0.0002	**0.9870** ± 0.0001	–	**2250** ± 75
L-GATr (ours)*	0.9423 ± 0.0002	**0.9870** ± 0.0001	**540** ± 20	2240 ± 70

Generation: Reconstructed events

- This task has the potential of **drastically cutting the computational cost of LHC simulations**. One of the main obstacles is the presence of sharp edges and long tails in high-dimensional distributions.
- We introduce the **first-ever Lorentz-equivariant generative model*,** based on an L-GATr vector field and trained with **Riemannian flow matching**.
- Our model improves the **generation of challenging feature distributions**, and it also outperforms all baselines across multiple process multiplicities.







*as far as we know