GLGENN: A Novel Parameter-Light Equivariant Neural Networks Architecture Based on Clifford Geometric Algebras

This work was prepared within the framework of the project 'Mirror Laboratories' HSE University ('Quaternions, geometric algebras, and applications').

Ekaterina Filimoshina[1], Dmitry Shirokov[1]
[1] HSE University, Moscow, Russia
filimoshinaek@gmail.com dm.shirokov@gmail.com

TL;DR: We propose a new equivariant neural network GLGENN that balances between expressiveness of geometric algebra-based models and parameter-efficiency.

Introduction & Motivation

The research focuses on **neural networks that are equivariant with respect to any pseudo-orthogonal transformation** (rotations, reflections, etc.).

These networks have **various applications in natural and computer sciences**, where tasks inherently involve equivariance to pseudo-orthogonal transformations (pseudo-orthogonal groups):

- modeling dynamical systems
- processing tasks involving point clouds
- motion capture
- particle physics
- analyzing molecular and protein properties
- estimating arterial wall-shear stress
- robotic planning
- ...

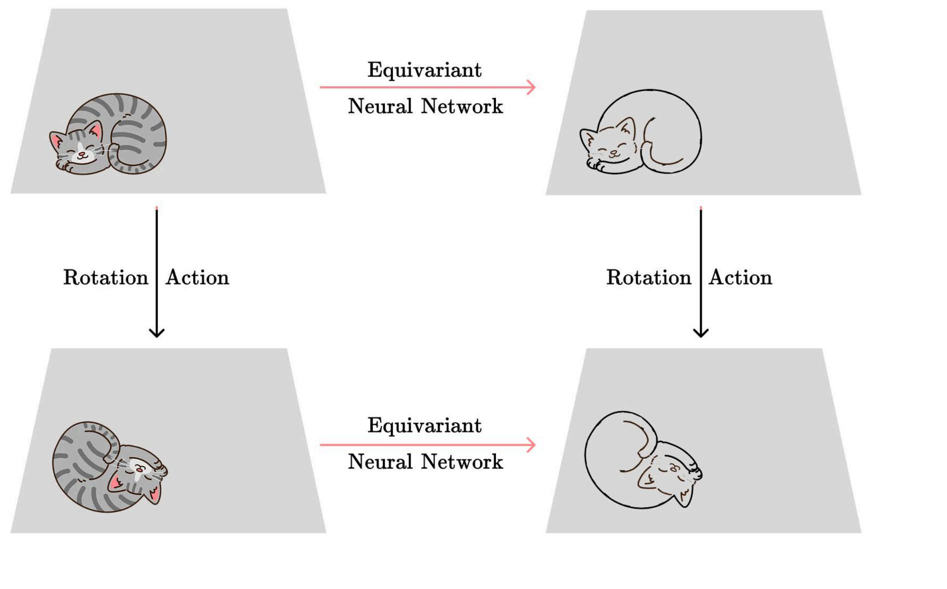

Figure 1. Example of how an equivariant neural network operates. In this example, the network detects edges of cats in images and is equivariant with respect to image rotations (a type of transformation from the pseudo-orthogonal group).

A neural network (function) L: X→Y is called **equivariant** iff it commutes with the actions $\circ_X$ and $\circ_Y$ of the group G onto the sets X and Y:

$$L(g \circ_X x) = g \circ_Y L(x), \quad \forall g \in G, \quad \forall x \in X.$$

In simple words: we get the same output if we transform the input to the neural network or transform the output.

SOTA Equivariant Models

Clifford Group Equivariant Neural Networks (CGENN); Clifford-Steerable Convolutional Neural Networks; Group Equivariant Convolutional Networks; Geometric Algebra Transformer; SE(3)-Transformers; Clifford Group Equivariant Simplicial Message Passing Networks; Equivariant Multilayer Perceptrons; ...

General Problem

Current equivariant neural networks, including SOTA models like CGENN, **suffer from parameter inefficiency.**

Excessive numbers of parameters increase the risk of **overfitting** and lead to **inefficient training times**, especially in case of small train datasets.

This limits their scalability and effectiveness in real-world applications requiring symmetry-preserving architectures.

Our Contribution

Introduction of Generalized Lipschitz Groups (GLG). They are applied in GLGENN and interesting for applications in theory of spin groups in physics and engineering.

Development and implementation of GLGENN architecture. GLGENN are adaptable to any pseudo-orthogonal groups equivariant task.

Equivariance proofs. New equivariant functions expand the toolkit for building equivariant neural networks.

Superior performance. GLGENN combine high performance with parameter efficiency, reducing overfitting and training time, while giving SOTA results on several benchmarking tasks.

In a Nutshell...

Aim: design, implement, test, and compare with competitors a new parameter-light architecture of equivariant neural networks, which we call Generalized Lipschitz Group Equivariant Neural Networks (GLGENN).

GLGENN — expressive power of CGENN / parameter efficiency

What we have done:

1. Introduce and study **generalized Lipschitz groups (GLG)** in geometric algebras.
2. Prove that **GLG are closely related to pseudo-orthogonal groups**: equivariance w.r.t. GLG ⟶ pseudo-orthogonal equivariance
3. Find mappings that are GLG-equivariant. Prove their equivariance.
   1. Projections onto the subspaces of geometric algebras determined by the grade involution and reversion.
   2. Polynomials of geometric algebra elements.
   3. Norm functions of geometric algebra elements.
4. Design **GLGENN layers** based on these mappings. Show efficiency of the proposed parameterization technique.
5. Implement and test GLGENN

GLGENN (Generalized Lipschitz Groups Equivariant Neural Networks)

GLGENN is a new architecture of neural networks **equivariant with respect to any pseudo-orthogonal transformation.**

Inputs and **outputs** are represented as **multivectors** (geometric algebra elements).

They encode geometric quantities such as **scalars**, **vectors**, oriented **areas (bivectors)** and **volumes (trivectors)**, and higher-dimensional objects (4-vectors, etc.):

A multivector has the form:

$$ue + u_1e_1 + \dots + u_ne_n + u_{12}e_{12} + \dots + u_{123}e_{123} + \dots + u_{1\dots n}e_{1\dots n}$$

(e.g., mass, temperature, or a one-hot-encoded feature; e.g., position or velocity; e.g., signed volume)

where $e, \dots, e_{1\dots n}$ are geometric algebra basis elements and $u, \dots, u_{1\dots n} \in \mathbb{R}$

GLGENN are **parameter-light**, since they operate in a unified manner across 4 fundamental subspaces of geometric algebras defined by the grade involution ($\widehat{\ }$) and reversion ($\widetilde{\ }$); they process geometric objects in groups with a step size of 4.

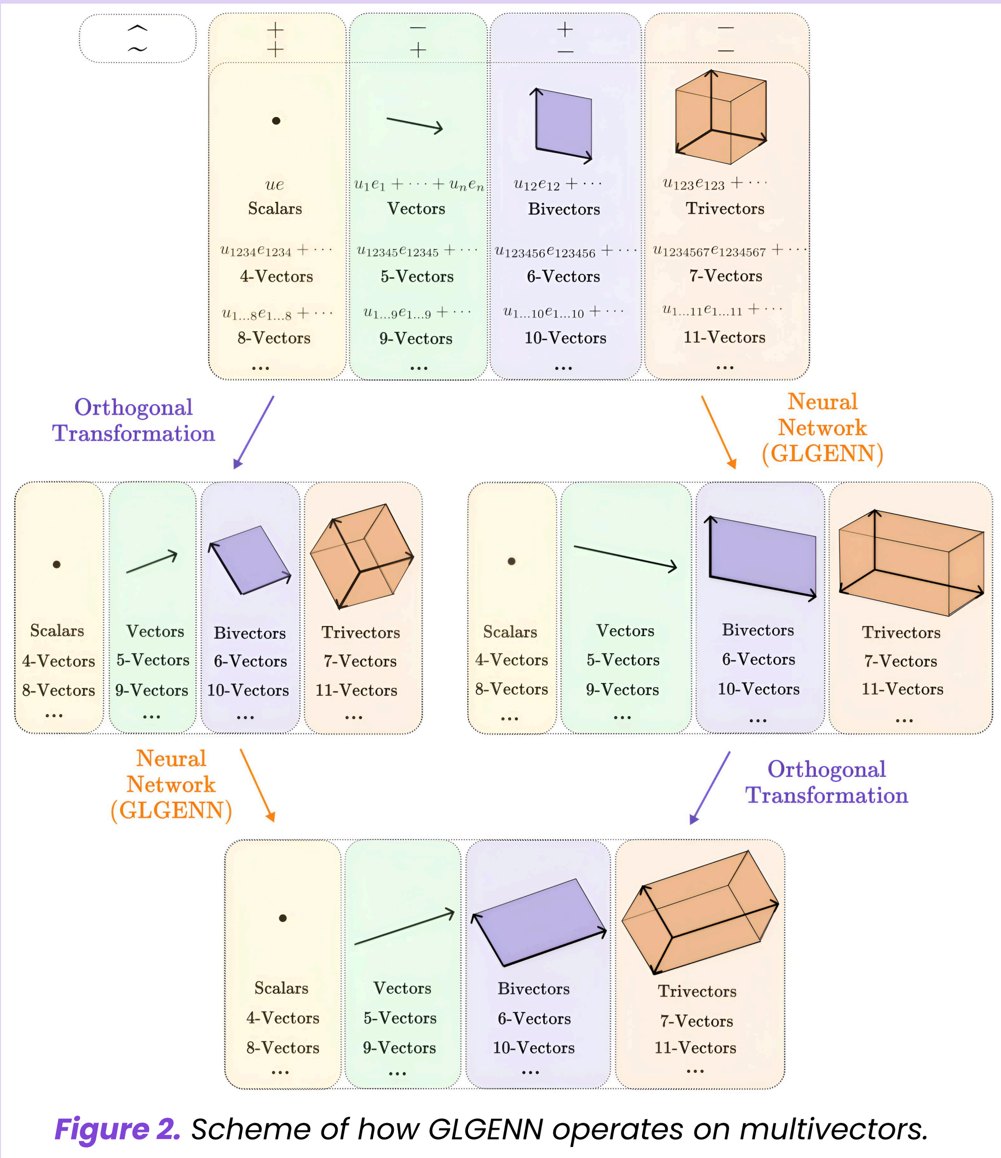

Figure 2. Scheme of how GLGENN operates on multivectors.

Geometric Algebra Linear Layer

Suppose $x_1, \dots, x_l \in C\ell$ are input data multivectors, where l is a number of input channels. A $C\ell^{\overline{k}}$-linear layer is constructed using

$$\langle y_{c_{out}} \rangle_{\overline{k}} := \sum_{c_{in}=0}^{l} \phi_{c_{out}c_{in}\overline{k}} \langle x_{c_{in}} \rangle_{\overline{k}},$$

(Polynomial; Projection)

where $y_{c_{out}} := \sum_{m=0}^{3} \langle y_{c_{out}} \rangle_{\overline{k}}$ is an output channel, $\phi_{c_{in}c_{out}\overline{k}} \in \mathbb{R}$ are optimizable coefficients, and c_{in} and c_{out} are used to denote the number of the input and output channels respectively.

Geometric Product Layer

A second order interaction term for the pair of multivectors x_1 and x_2 has the form $\langle \langle x_1 \rangle_{\overline{i}} \langle x_2 \rangle_{\overline{j}} \rangle_{\overline{k}}$, where $i, j, k = 0, 1, 2, 3$. All terms from the subspace $C\ell^{k}$, $k = 0, 1, 2, 3$, resulting from the interaction of x_1 and x_2 are parameterized as

$$P(x_1, x_2)^{\overline{k}} := \sum_{i=0}^{3} \sum_{j=0}^{3} \phi_{ijk} \langle \langle x_1 \rangle_{\overline{i}} \langle x_2 \rangle_{\overline{j}} \rangle_{\overline{k}},$$

(Polynomial; Projections)

where $\phi_{ijk} \in \mathbb{R}$ are optimizable parameters.

Geometric Algebra Normalization

For numerical stability of $C\ell^{\overline{k}}$-geometric product layers, we apply normalization to the four projections $\langle x_{c_{in}} \rangle_{\overline{0}}, \langle x_{c_{in}} \rangle_{\overline{1}}, \langle x_{c_{in}} \rangle_{\overline{2}}, \langle x_{c_{in}} \rangle_{\overline{3}}$ for each multivector $x_{c_{in}}, c_{in} = 1, \dots, l$:

$$\langle x_{c_{in}} \rangle_{\overline{k}} \mapsto \frac{\langle x_{c_{in}} \rangle_{\overline{k}}}{\sigma(\phi_{c_{in}\overline{k}})(\langle \widetilde{\langle x_{c_{in}} \rangle_{\overline{k}}} \langle x_{c_{in}} \rangle_{\overline{k}} \rangle_0 - 1) + 1},$$

(Polynomial; Projections; Norm function)

where $\sigma(x) := \frac{1}{1+e^{-x}} \in (0,1)$ is the logistic sigmoid function and $\phi_{c_{in}\overline{k}} \in \mathbb{R}$ are optimizable parameters.

Conjugation Operations Layer

These layers are based on the concept of conjugation operations in geometric algebras. For an input data multivector $x_{c_{in}}, c_{in} = 1, \dots, l$, a conjugation operations layer is constructed using

$$x_{c_{in}} \mapsto \sum_{k=0}^{n} \phi_{c_{in}k} \langle x_{c_{in}} \rangle_k, \quad \phi_{c_{in}k} = \pm 1,$$

where $\phi_{c_{in}k} \in \{-1, 1\}$ are optimizable parameters.

Geometric Algebras

Let us consider **geometric (Clifford) algebras** $C\ell(V) = C\ell_{p,q,r}$, $p+q+r = n \geq 1$, over a vector space V with a symmetric bilinear form, where V can be real $\mathbb{R}^{p,q,r}$ or complex $\mathbb{C}^{p+q,0,r}$. The identity element of $C\ell$ is denoted by $e \equiv 1$, the generators are denoted by $e_a, a = 1, \dots, n$, and satisfy

$$e_a e_b + e_b e_a = 2\eta_{ab} e, \qquad \forall a, b = 1, \dots, n,$$

where $\eta = (\eta_{ab})$ is the diagonal matrix with p times +1, q times -1, and r times 0 on the diagonal in the real case $C\ell(\mathbb{R}^{p,q,r})$ and $p+q$ times +1 and r times 0 on the diagonal in the complex case $C\ell(\mathbb{C}^{p+q,0,r})$.

Grading

Consider the **subspaces of fixed grades** $C\ell^k$, $k = 0, \dots, n$. Their elements are linear combinations of basis elements $e_{a_1 \dots a_k} := e_{a_1} \cdots e_{a_k}$, $a_1 < \dots < a_k$. Any element $U \in C\ell$ (**multivector**) can be represented as the sum

$$U = \langle U \rangle_0 + \dots + \langle U \rangle_n, \quad \langle U \rangle_k \in C\ell^k, \quad k = 0, \dots, n.$$

Consider such conjugation operations as **grade involution** and **reversion** defined on an arbitrary $U \in C\ell$ as

$$\widehat{U} := \sum_{k=0}^{n} (-1)^k \langle U \rangle_k, \quad \widetilde{U} := \sum_{k=0}^{n} (-1)^{\frac{k(k-1)}{2}} \langle U \rangle_k.$$

The grade involution and reversion define four **subspaces of quaternion types**: $C\ell^{\overline{k}} = \{U \in C\ell : \ \widehat{U} = (-1)^k U, \ \widetilde{U} = (-1)^{\frac{k(k-1)}{2}} U\}$ for $k = 0, 1, 2, 3$. I.e. $C\ell^{\overline{k}} = C\ell^k \oplus C\ell^{k+4} \oplus C\ell^{k+8} \oplus \cdots$

Any element $U \in C\ell$ can be represented as a sum

$$U = \langle U \rangle_{\overline{0}} + \langle U \rangle_{\overline{1}} + \langle U \rangle_{\overline{2}} + \langle U \rangle_{\overline{3}}, \quad \langle U \rangle_{\overline{m}} \in C\ell^{\overline{m}}, \ m = 0, 1, 2, 3.$$

Lipschitz Groups and Orthogonal Groups

Ordinary Lipschitz group: $\tilde{\Gamma}^1 := \{T \in C\ell^\times : \ \hat{T} C\ell^1 T^{-1} \subseteq C\ell^1\}$

Generalized Lipschitz group: $\tilde{\Gamma}^{\overline{1}} := \{T \in C\ell^\times : \ \hat{T} C\ell^{\overline{1}} T^{-1} \subseteq C\ell^{\overline{1}}\}$

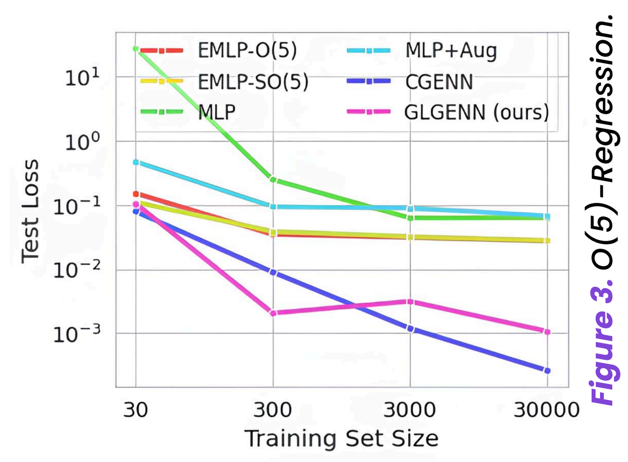

Figure 3. O(5)-Regression.

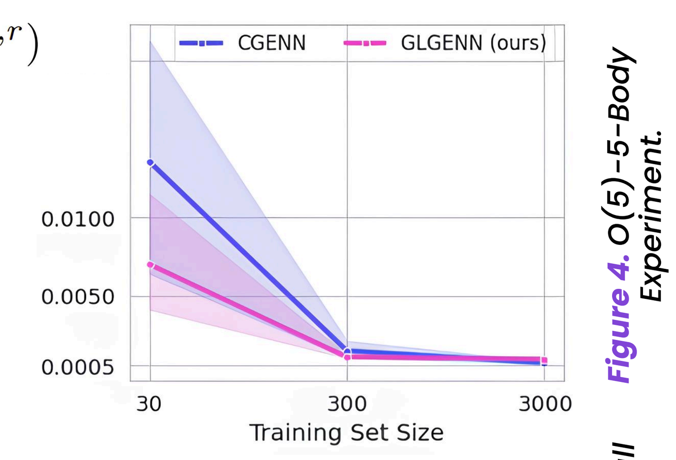

Figure 4. O(5)-5-Body Experiment.

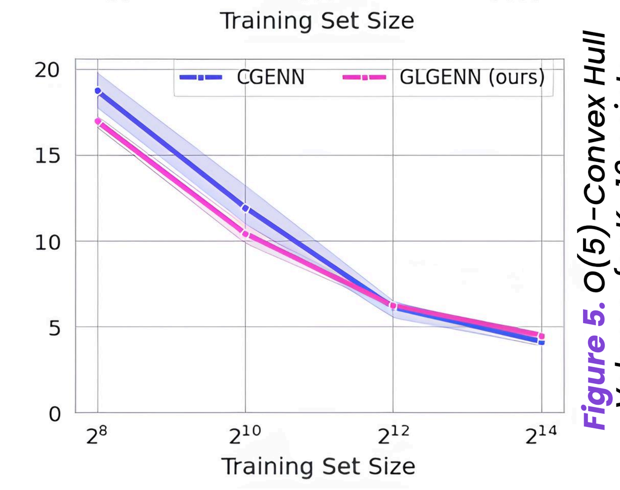

Figure 5. O(5)-Convex Hull Volume, for K=16 points.

We have: $\tilde{\Gamma}^{\overline{1}} / \mathbb{R}^\times \cong \mathrm{O}(p,q)$ where $\mathrm{O}(p,q) = \{A \in \mathrm{Mat}(n, \mathbb{R}) : A^{\mathrm{T}} \eta A = \eta\}$

Experiments

O(5)-Regression

Task: Estimate value of the equivariant function $\sin(\|x_1\|) - \|x_2\|^3/2 + \frac{x_1^T x_2}{\|x_1\|\|x_2\|}$, where $x_1, x_2 \in \mathbb{R}^{5,0}$ are vectors sampled from a Gaussian distribution.

O(5)-N-body Experiment

Task: Given masses, initial positions, and velocities of N charged particles (bodies) in $\mathbb{R}^{5,0}$, predict their final positions after the system evolves.

Architecture: We construct a graph neural network based on the message passing. Bodies are nodes, their pairwise interactions are edges. The message and update networks are GLGENN.

O(5)-Convex Hull Volume Estimation

Task: Estimate the volume of a convex hull generated by K=16, 256, or 512 points in $\mathbb{R}^{5,0}$

Architecture: For GLGENN, we adopt a similar architecture to CGENN but utilize our $C\ell^{\overline{k}}$-linear and geometric product layers instead of ordinary linear and geometric product layers.

References

Paper

Filimoshina E., Shirokov D. **GLGENN: A Novel Parameter-Light Equivariant Neural Networks Architecture Based on Clifford Geometric Algebras.** Proceedings of the 42nd International Conference on Machine Learning, ICML 2025, Vancouver, Canada, July 13-19, 2025

Code

https://github.com/katyafilimoshina/glgenn

Theoretical Results

Filimoshina, E. and Shirokov, D. **On generalization of Lipschitz groups and spin groups.** Mathematical Methods in the Applied Sciences, 47(3):1375–1400, 2024.

Shirokov, D. **On inner automorphisms preserving fixed subspaces of Clifford algebras.** Advances in Applied Clifford Algebras, 31:30, 2021.

Filimoshina, E. and Shirokov, D. **On some Lie groups in degenerate Clifford geometric algebras.** Advances in Applied Clifford Algebras, 33(44), 2023.

Filimoshina, E. and Shirokov, D. **Generalized degenerate Clifford and Lipschitz groups in geometric algebras.** Advances in Applied Clifford Algebras, 35(29), 2025.