Generalized Geometric Algebra Transformer

Yuxin Yao, Christian Hockey, Joan Lasenby
CUED, Cambridge

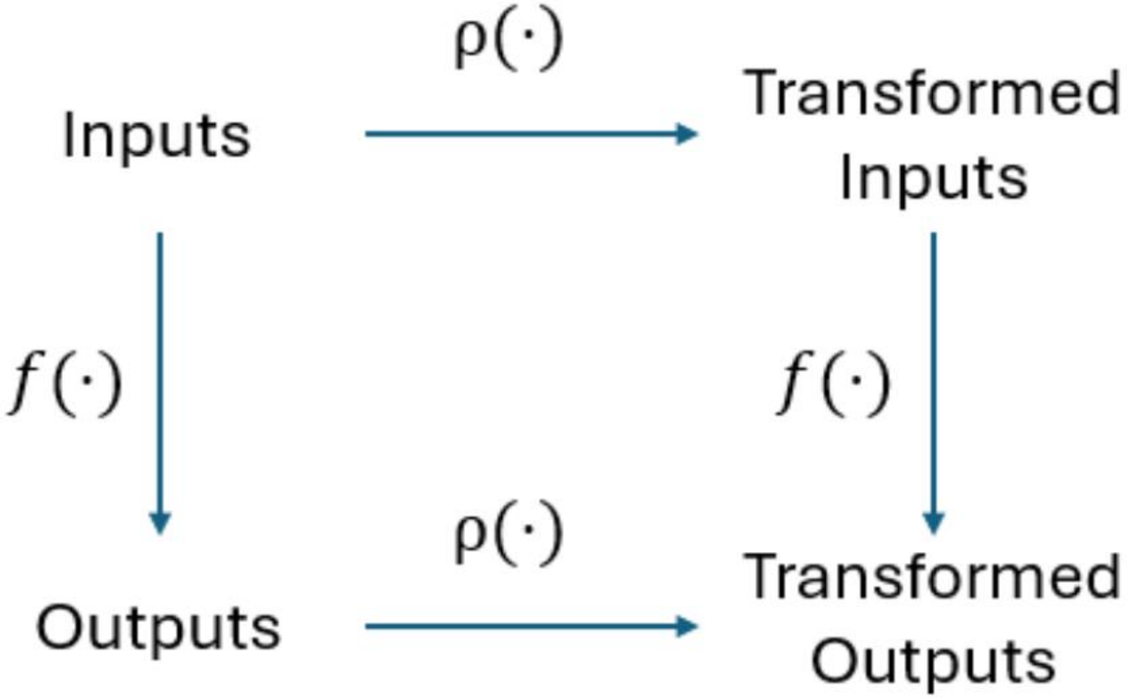
Context

- Latest Geometric Algebra Transformer : GATr
  - First Transformer model in GA, with great potential in understanding the 3D geometry of input data.
  - Fixed GA signature
- Structure of Generalized Geometric Algebra Transformer (GGATr)
- Experiments
  - Camera Pose Estimation
  - N-Body Dynamics
  - Protein Structure Prediction

Ref: Brehmer, J., De Haan, P., Behrends, S., and Cohen, T. S. (2024). Geometric algebra transformer. Advances in Neural Information Processing Systems, 36.
Ruhe, D., Brandstetter, J., and Forré, P. (2024). Clifford group equivariant neural networks. Advances in Neural Information Processing Systems, 36.

Equivariance

$$f(\rho(x)) = \rho(f(x))$$



Ref: Brehmer, J., De Haan, P., Behrends, S., and Cohen, T. S. (2024). Geometric algebra transformer. Advances in Neural Information Processing Systems, 36.
Ruhe, D., Brandstetter, J., and Forré, P. (2024). Clifford group equivariant neural networks. Advances in Neural Information Processing Systems, 36.

Structure of GA Layers

- Linear Layer
- Non-linear Layer
- Sandwich product layer
- Layer Norm
- Geometric Product Layer
- Join Layer

Structure of GA Layers

- Linear Layer $\langle y_j \rangle_r = \sum_{i=1}^{N} K_{ijr} \langle x_i \rangle_r$
  - ➢ Equivariance proof: 1. $R\langle x \rangle_r \tilde{R} = \langle R x \tilde{R} \rangle_r$

    2. $R\left(\sum_{i=1}^{N} a_i x_i\right)\tilde{R} = \sum_{i=1}^{N} a_i R x_i \tilde{R}$

- Non-linear Layer $y_i = \phi(h(x_i))x_i$ where $h(x) = \langle x\tilde{x} \rangle_0$

Structure of GA Layers

- Layer Norm

$$y_i = \frac{x_i}{\sum_{k=1}^{N} \sqrt{|\langle x_k \tilde{x}_k \rangle_0|}/N}$$

Dimension of x

- Sandwich Product Layer (Not equivariant)

$$y_j = \sum_{i=1}^{N} w_{ij} x_i w_{ij}^{-1}$$

Multivector weights

Structure of GA Layers

- Geometric Product Layer

$$y_i = \text{Linear}_1(x)_i \text{Linear}_2(x)_i$$

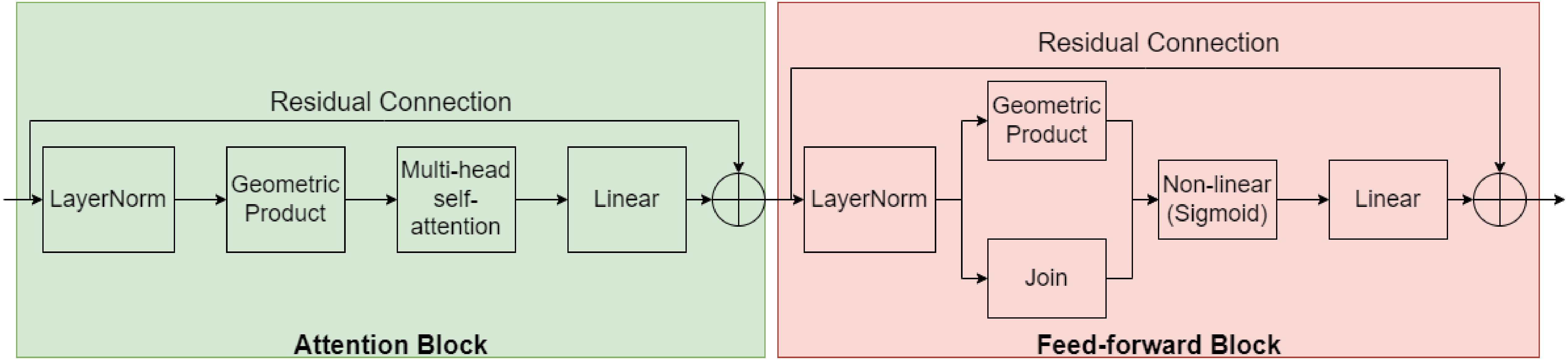
- Join/Meet Layer

$$y_i = \text{Linear}_1(x)_i \vee \text{Linear}_2(x)_i$$

  - define Join/meet: $u \vee v = (u^* \wedge v^*)^* = \left(\left(uI^{-1}\right) \wedge \left(vI^{-1}\right)\right) I^{-1}$
  - Finding the dual by. $x \wedge x^* = I$ If exist basis square to 0
    $x^* = xI$ Otherwise

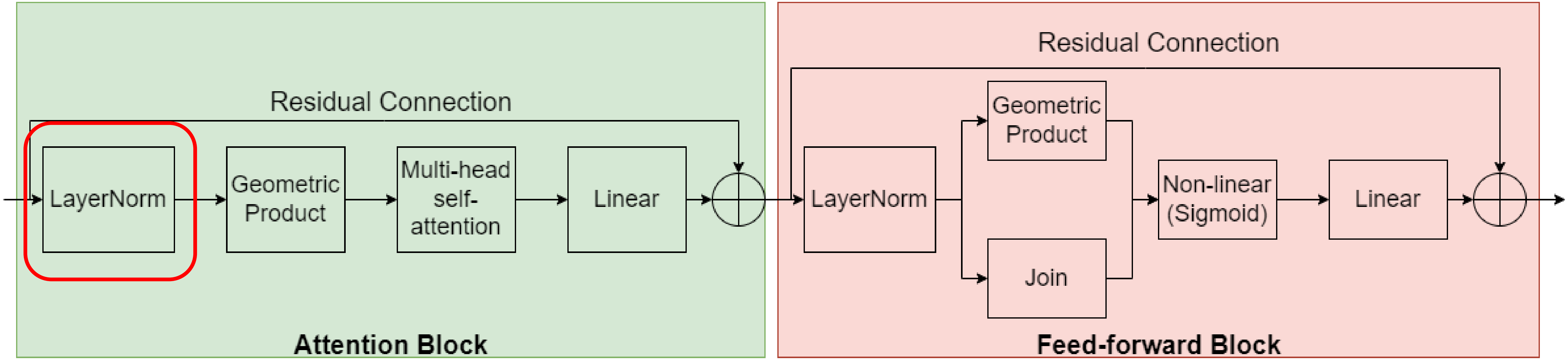
UNIVERSITY OF CAMBRIDGE

Structure of GGATr

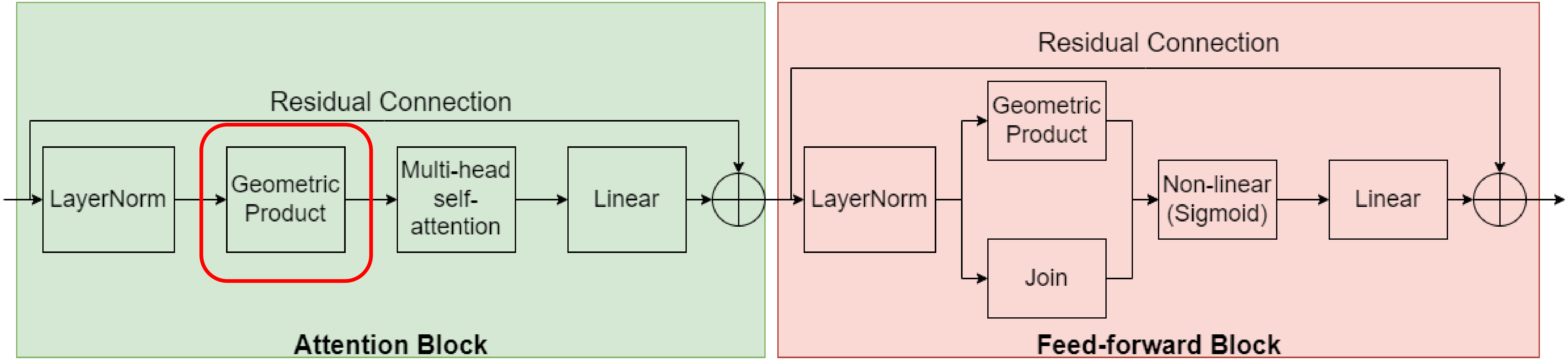


UNIVERSITY OF CAMBRIDGE

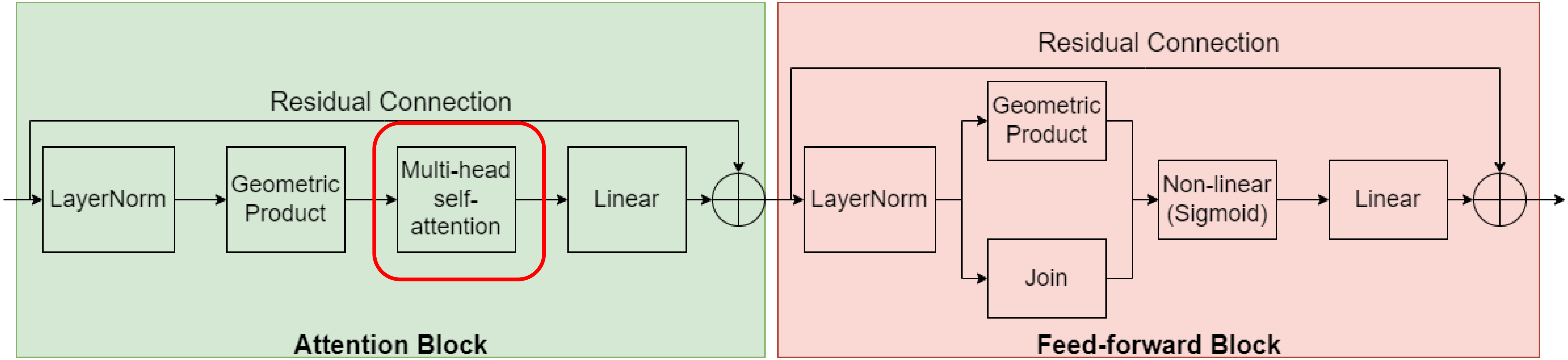
Structure of GGATr



Structure of GGATr



Structure of GGATr

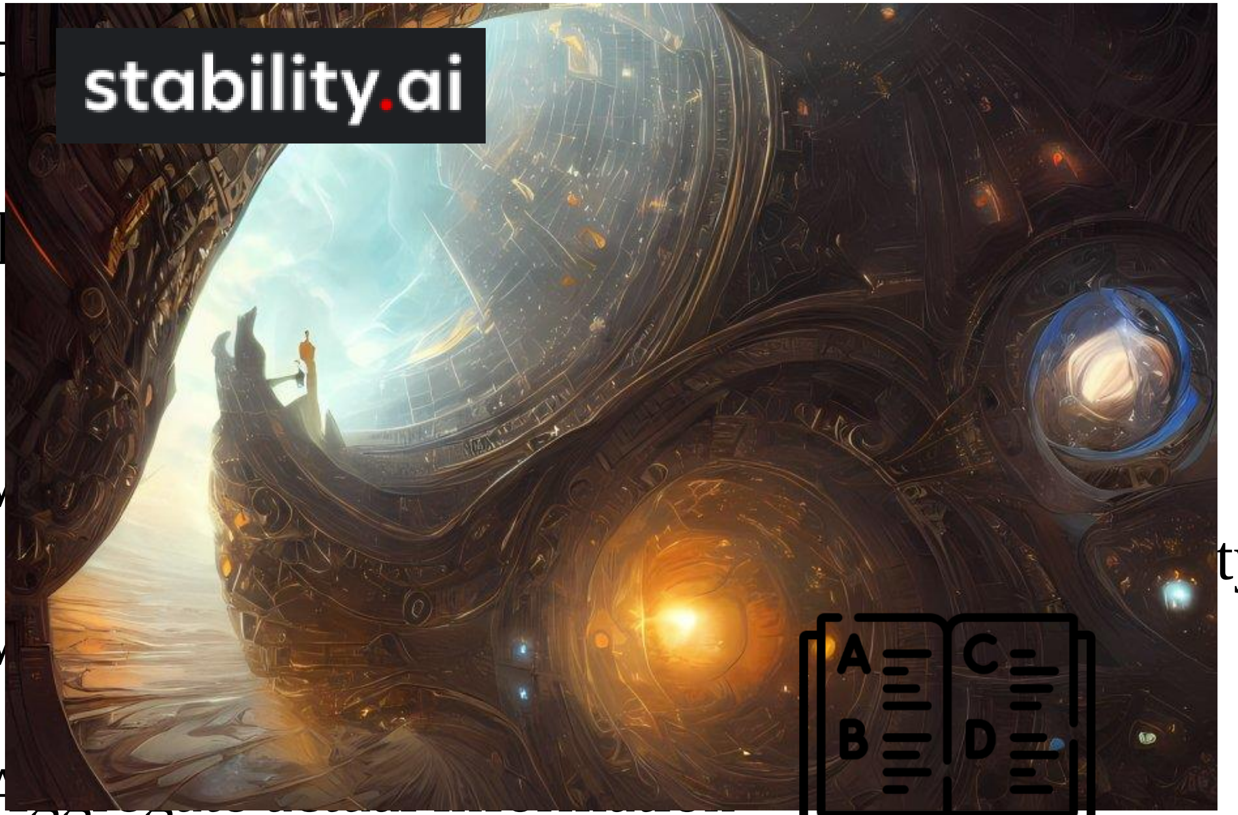


UNIVERSITY OF CAMBRIDGE

General Attention Mechanism

- Self-attention: Focus on different part

ChatGPT



ty

Value

Aggregate actual information I hold.

Attention Head

- Our Attention Mechanism

$$y_i = \text{Softmax}\left(\frac{\langle q_i \cdot k_i \rangle_0}{\sqrt{n_b}}\right) v_i$$

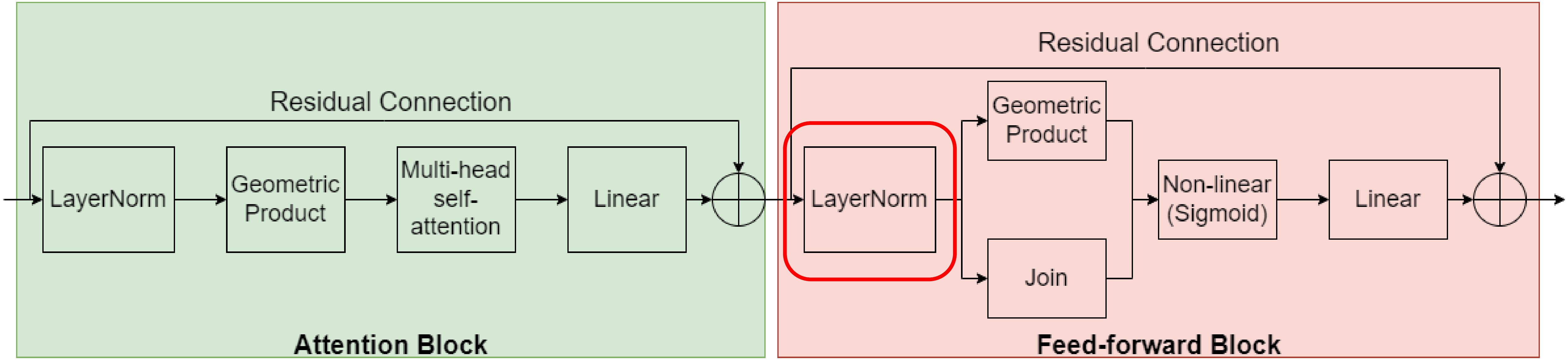
Number of basis

- QualComm's GATr: Distance-aware dot product attention

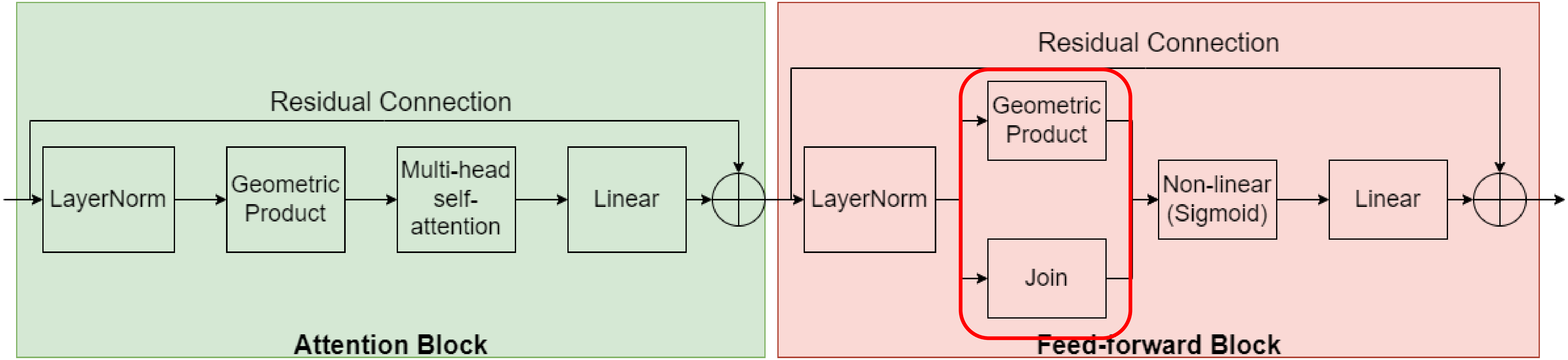
$$y_i = \text{Softmax}\left(\frac{\sum \langle q_i \cdot k_i \rangle}{\sqrt{8 n_i}}\right) v_i$$

Number of basis
not squares to 0

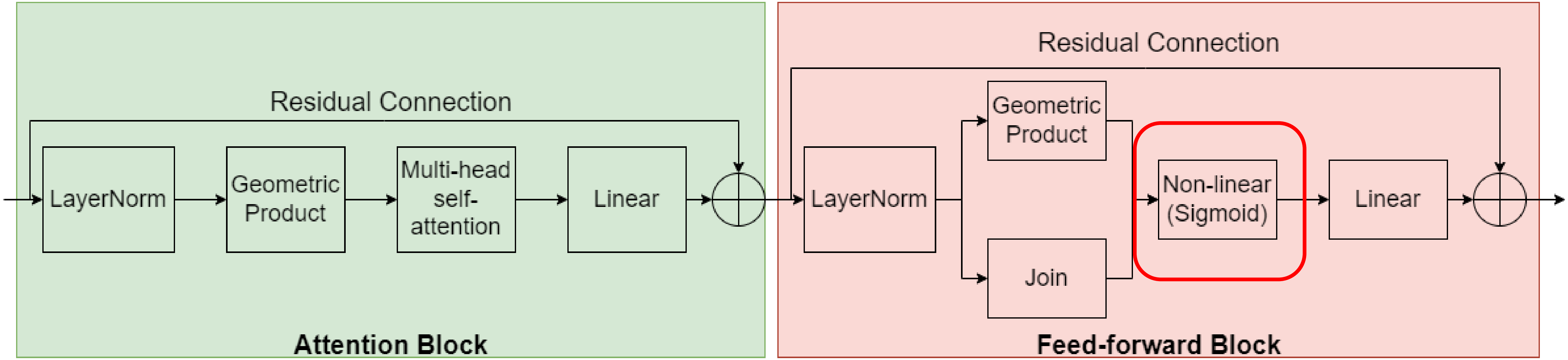
Structure of GGATr – Feed-forward Block



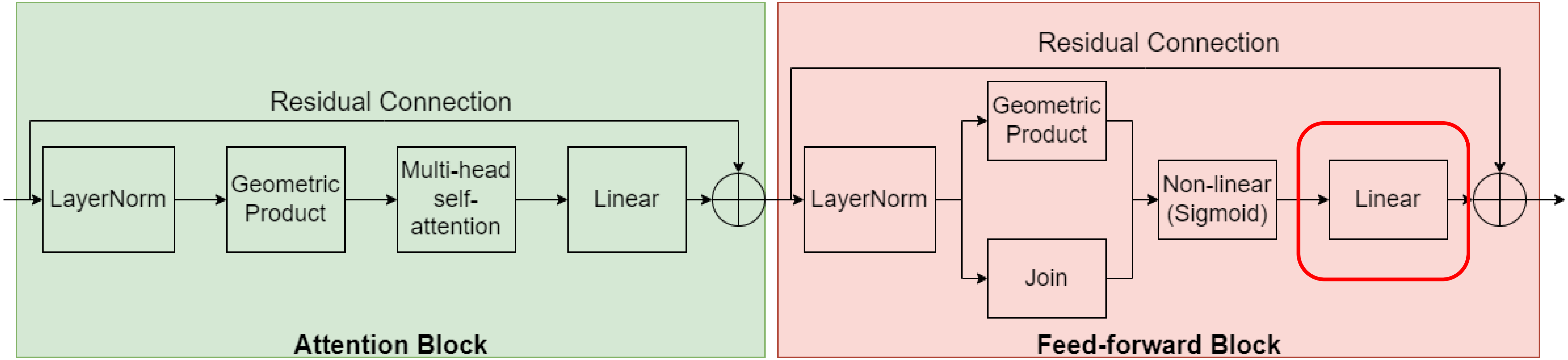
Structure of GGATr – Feed-forward Block



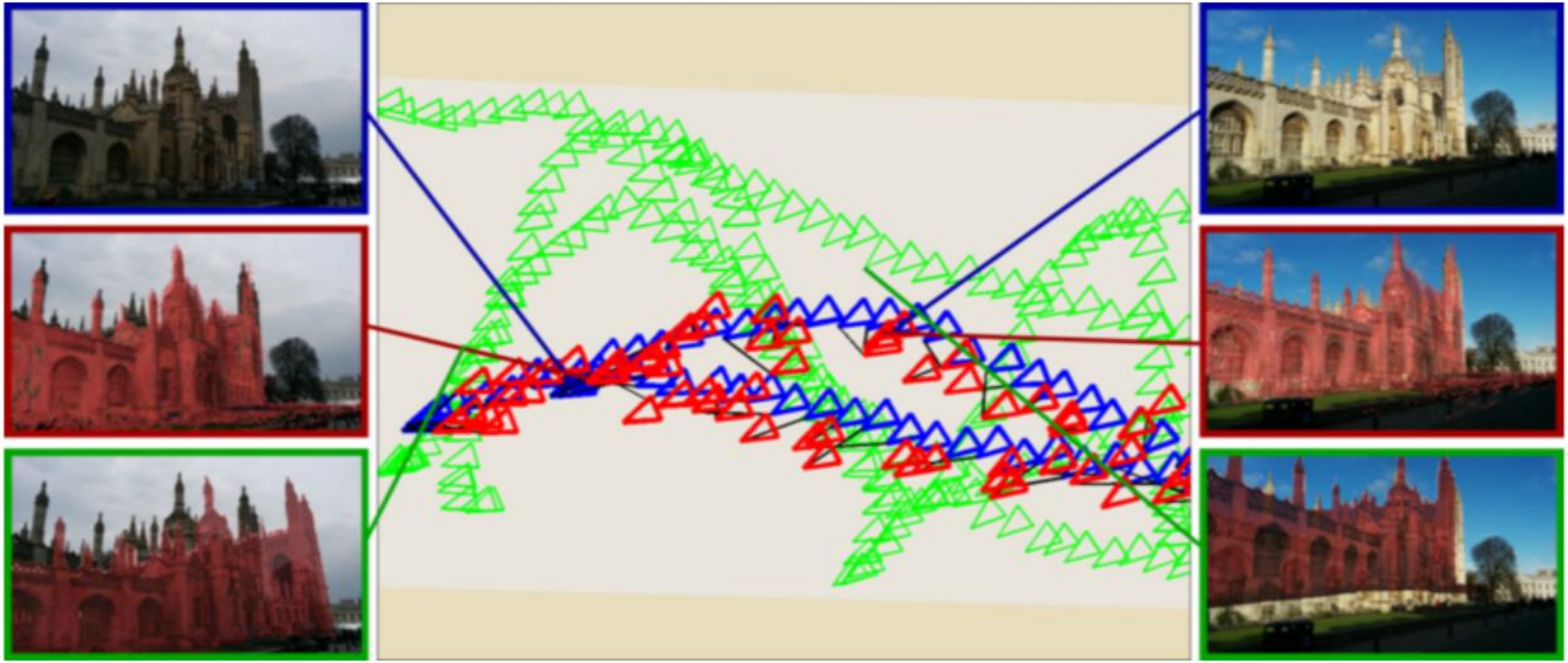
Structure of GGATr – Feed-forward Block



Structure of GGATr – Feed-forward Block

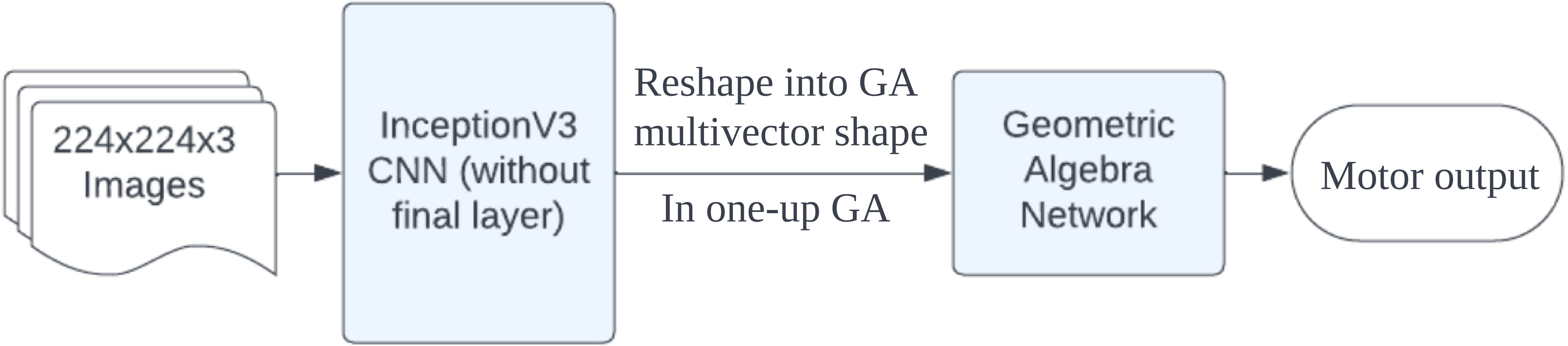


Experiments – PoseNet



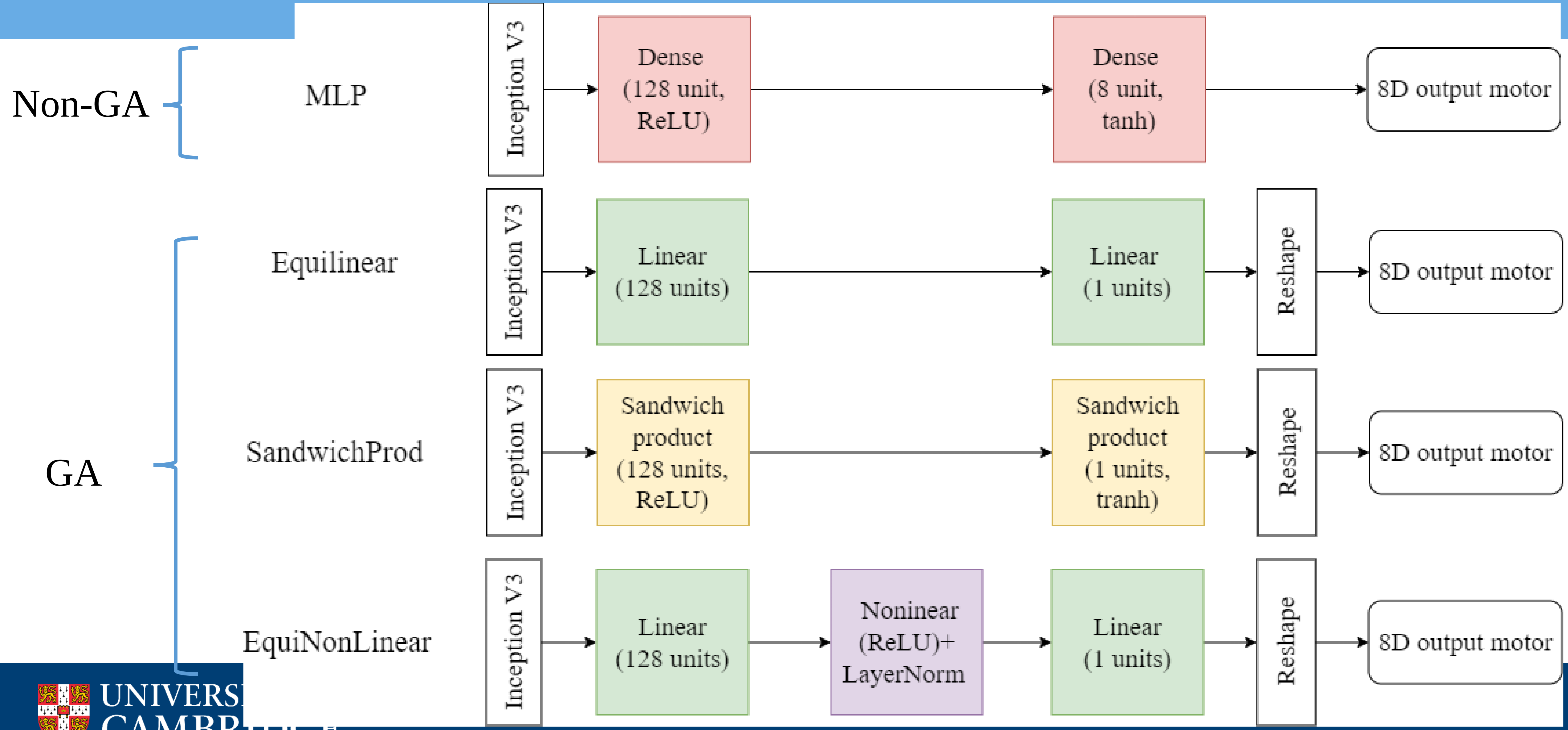
Ref: Kendall, A., Grimes, M., and Cipolla, R. (2015). Posenet: A convolutional network for real-time 6-dof camera relocalization. In Proceedings of the IEEE international conference on computer vision, pages 2938–2946.

Experiments – PoseNet



Ref: Pepe, A., Lasenby, J., and Buchholz, S. (2024). Cgaposenet+ gcan: A geometric clifford algebra network for geometry-aware camera pose regression. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pages 6593–6603

Posenet – Network Structure



PoseNet - Evaluation

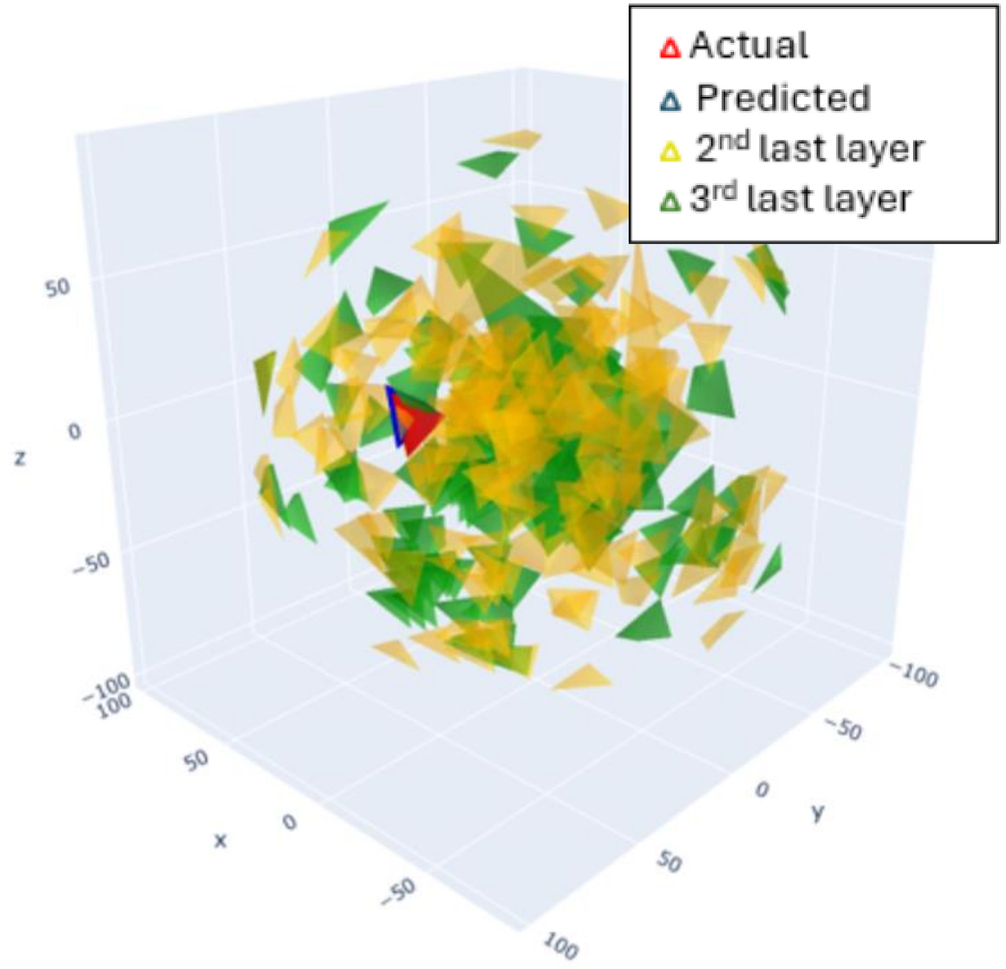
Model	No. of parameters	Training Loss (MSE)	Validation Loss (MSE)
MLP(non-GA)	263304	8.03×10^{-4}	0.0185
EquiLinear	164609	2.77×10^{-4}	0.0106
SandwichProduct	263297	3.47×10^{-4}	0.0112
EquiNonLinear	165249	2.90×10^{-4}	0.0083

- All networks with rotor product layer perform better.
- Non-linearity further improve the performances.

PoseNet - Visualization

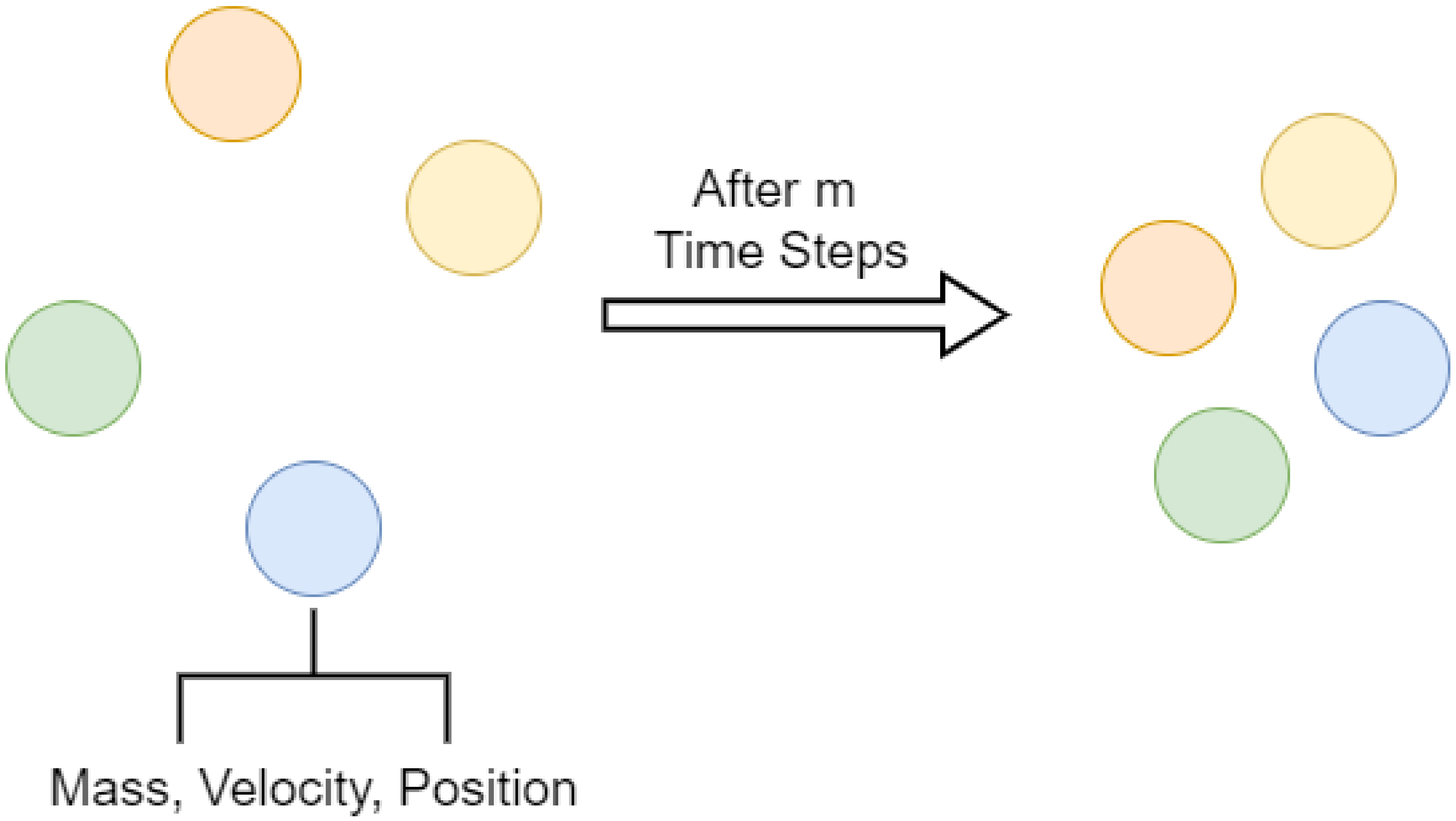
- Visualize the convergence of camera poses

- Intermediate tensors are motors



Ref: Pepe, A., Lasenby, J., and Buchholz, S. (2024). Cgaposenet+ gcan: A geometric clifford algebra network for geometry-aware camera pose regression. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pages 6593–6603

N-body Dynamics Problem

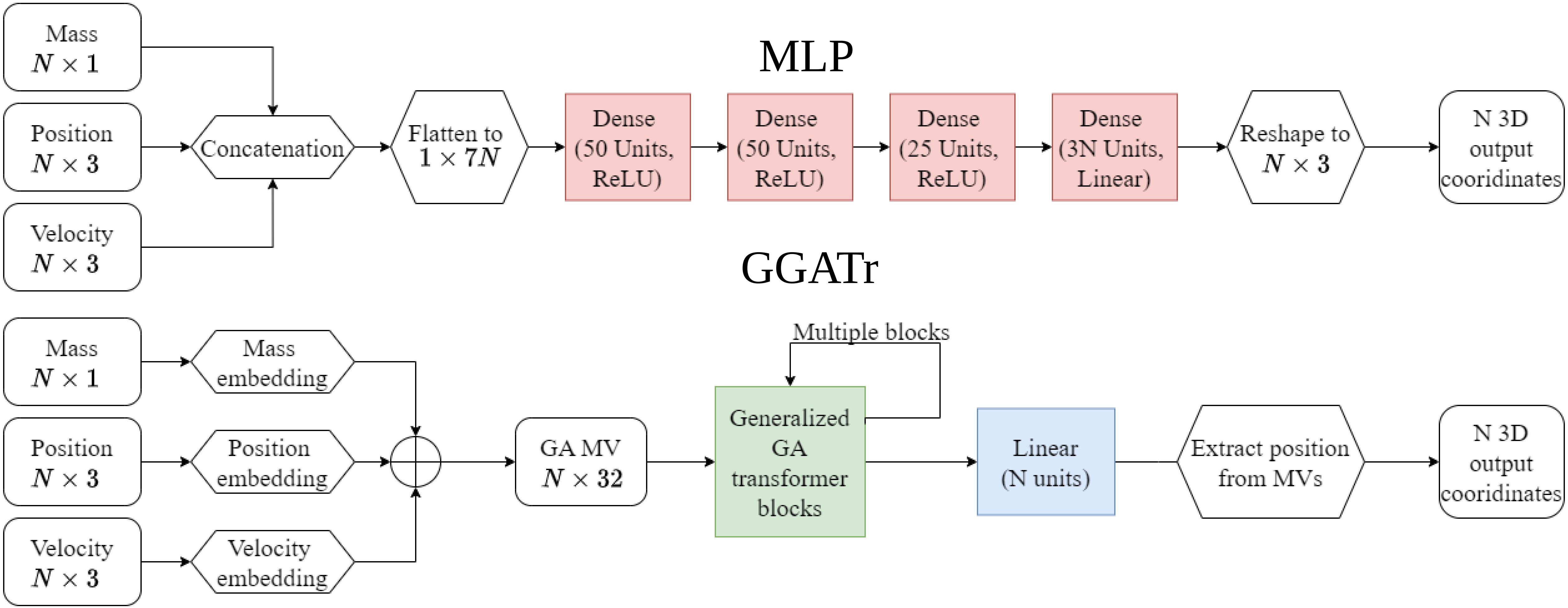


N-body Dynamics - Embedding

- In CGA, we define $n = e_+ + e_-$ and $\overline{n} = e_+ - e_-$
- Mass: Scalar
- Position: $X = x + \frac{1}{2}x^2 n - \frac{1}{2}\overline{n}$
- Velocity: $V = (v + (x \cdot v)n)I^{-1}$

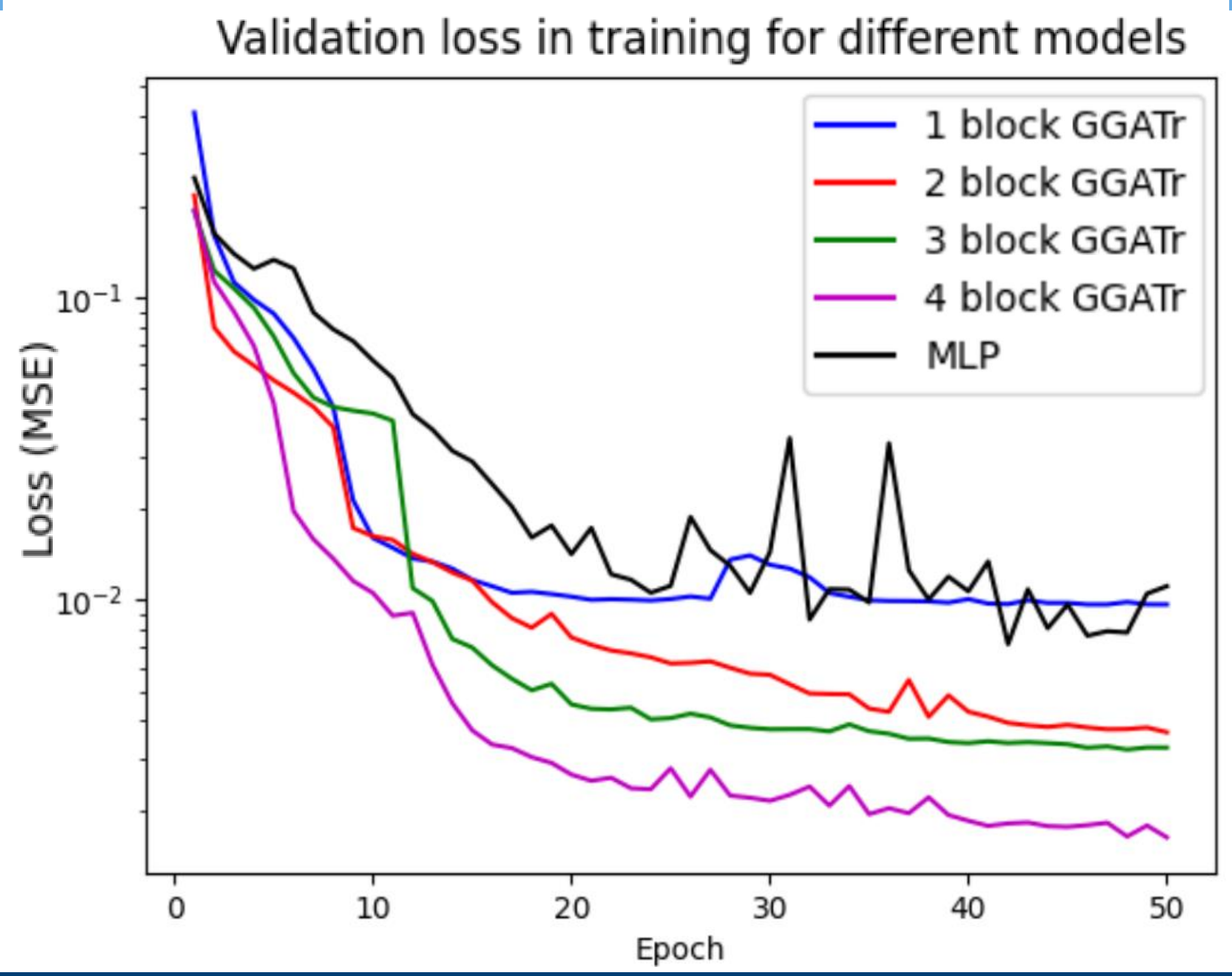
UNIVERSITY OF CAMBRIDGE

N-body Dynamics

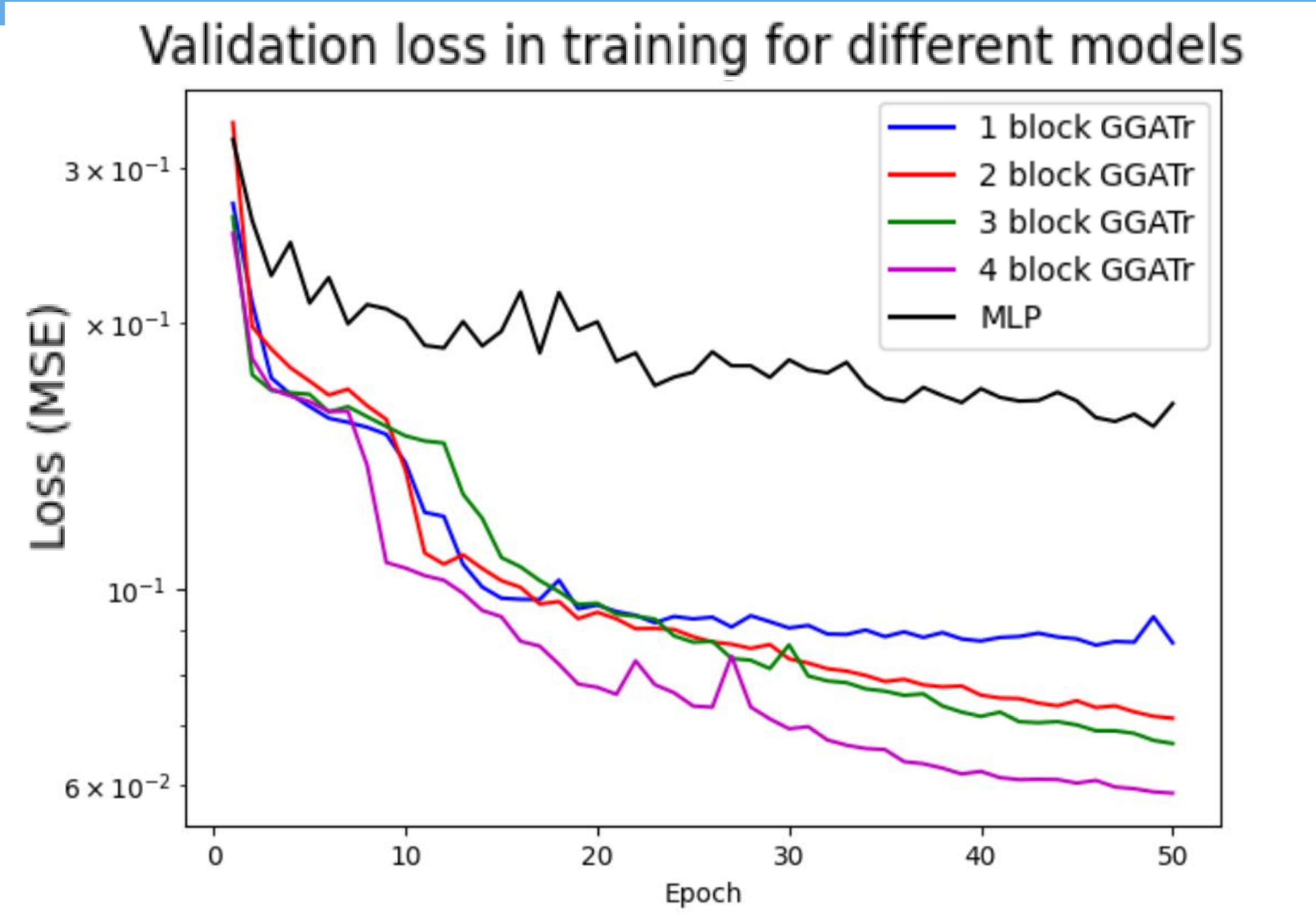


N-body Dynamics – Evaluation – 0.1s gap

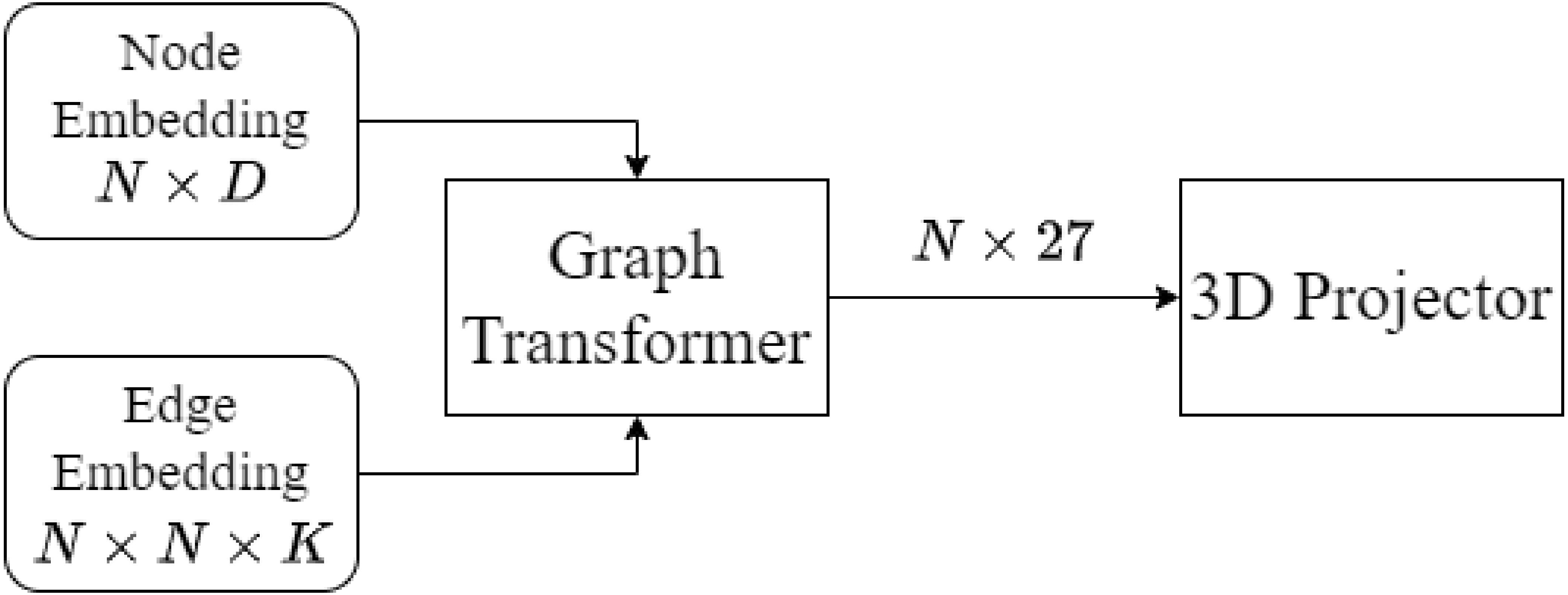
Model	No. of parameters	Validation Loss (MSE)
MLP	5587	0.0111
GGA 1 block	1496	0.0097
GGA 2 block	2892	0.0037
GGA 3 block	4288	0.0033
GGA 4 block	5684	0.0016
Qualcomm's GATR	1,900,000	Approx. 10^{-4}



N-body Dynamics – 1s gap



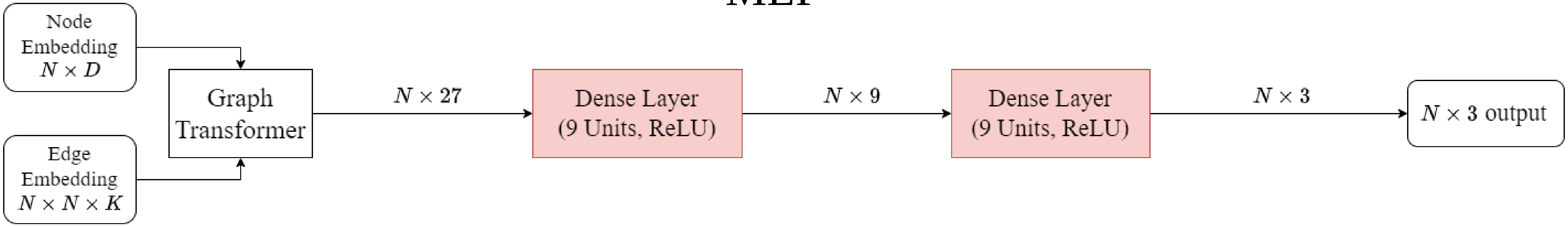
Protein Structure Prediction



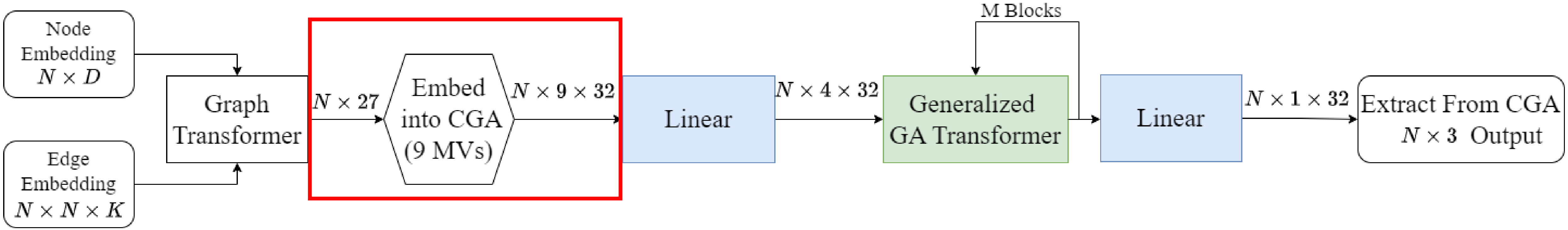
Pepe, Alberto, Sven Buchholz, and Joan Lasenby. "Clifford Group Equivariant Neural Network Layers for Protein Structure Prediction." Northern Lights Deep Learning Conference. PMLR, 2024.

Protein Structure Prediction - Structure

MLP



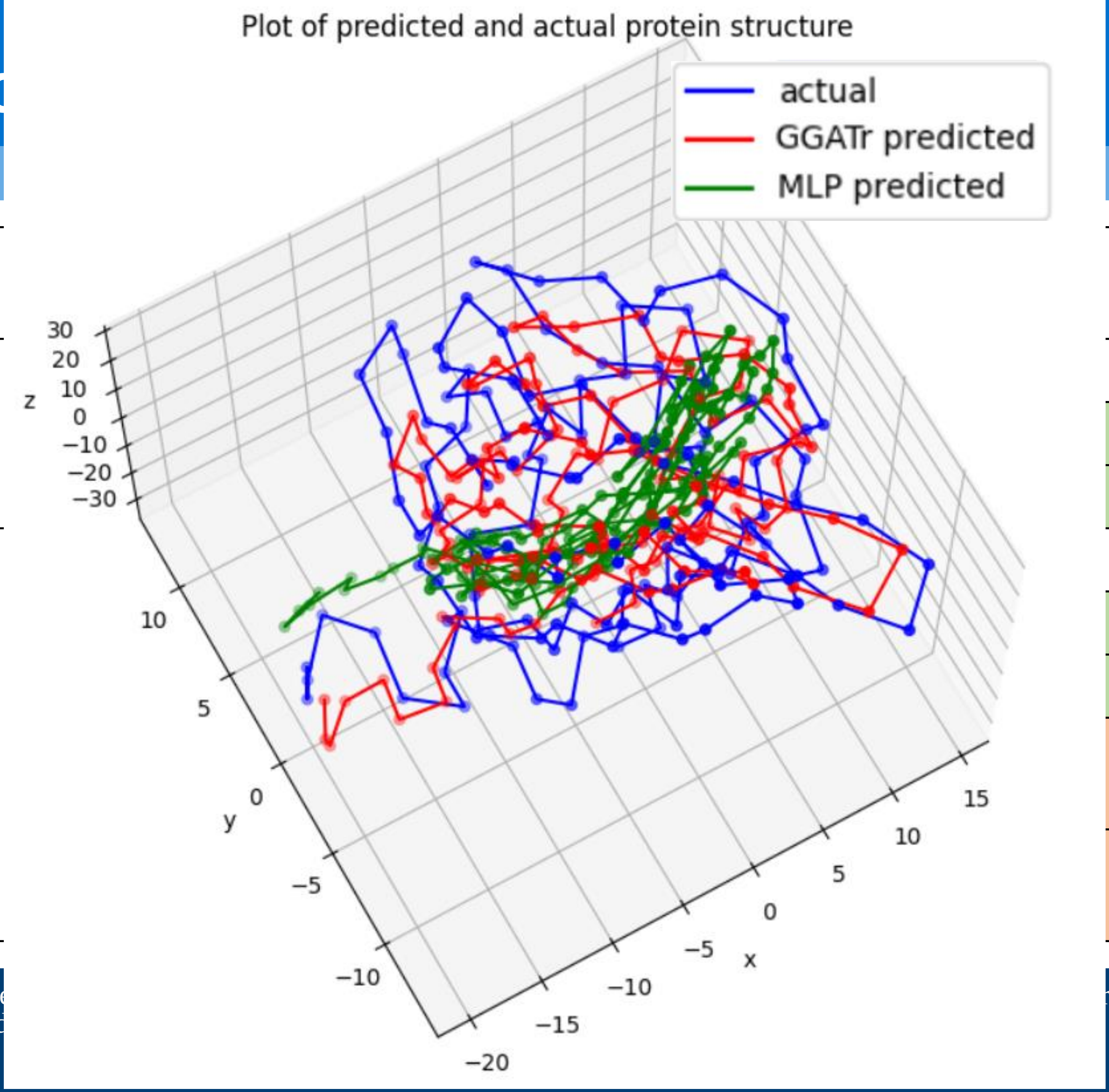
GGATr



Ref: Pepe, Alberto, Sven Buchholz, and Joan Lasenby. "Clifford Group Equivariant Neural Network Layers for Protein Structure Prediction." Northern Lights Deep Learning Conference. PMLR, 2024.

Protein Struc

Dataset
Validation
Test



Conclusion

- Constructed a Generalized Geometric Algebra Transformer.
- Demonstrated the impact of GA networks even used with an non-equivariant backbone.
- The GA equivariant layer outperforms both non-geometric methods and non-equivariant methods.

Thank you!

Q&A – Equivariance Proof

- Non-linear Layer
  - Function h is invariant to versor sandwich product -> constant scalar for given x $h(x) = h(Rx\tilde{R})$
  - Gated nonlinearity follows the equivariance of multiplication by scalar $h(x)$
  - Involve all element of a grade in nonlinearty but instable

Q&A – Equivariance Proof

- Join Layer
  - Equivariance of Dual

$$Rx I^{-1} \tilde{R} = Rx \tilde{R} R I^{-1} \tilde{R} = Rx \tilde{R} I^{-1}$$

  - Pesudoscalar invariance.
  - Define outer product:

$$A_r \wedge B_s = \langle A_r B_s \rangle_{r+s}$$

  - Express multivectors to sum of n+1 multivectors of a single grades.
  - By linearity of outer product, and equivariance of grade projection.
  - Dual, antidual and outer product are equivariant, so Join Layer equivariant.

Q&A – Equivariance Proof

- Self-Attention Layer

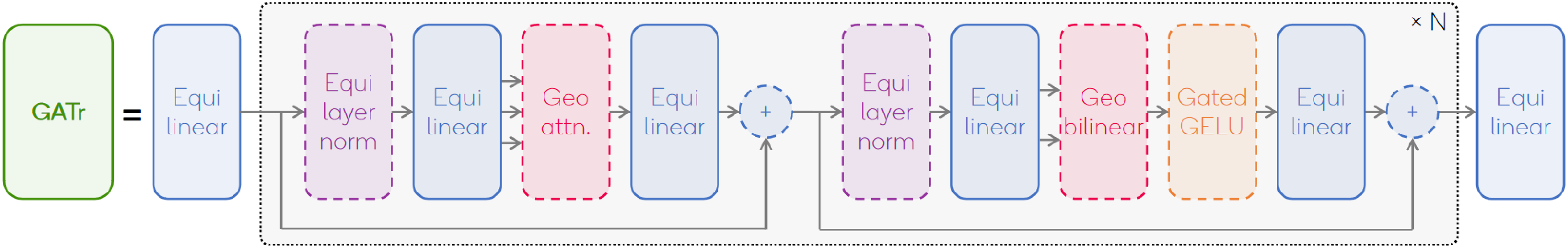
$$y_i = \mathrm{Softmax}\left(\frac{\langle q_i \cdot k_i \rangle_0}{\sqrt{n_b}}\right) v_i$$

- Use Same logic form $y_i = \phi(h(x_i))x_i$

- Invariant function: 0-grade projection of inner product

$$\langle Rx\tilde{R} \cdot Ry\tilde{R} \rangle_0 = \langle x \cdot y \rangle_0$$

Q&A – QualComm's GATr



- LayerNorm: $x / \sqrt{\mathbb{E}_c \langle x, x \rangle}$ GatedGeLU
- GeoBilinear: EquiJoin & GP layer

Ref: Brehmer, J., De Haan, P., Behrends, S., and Cohen, T. S. (2024). Geometric algebra transformer. Advances in Neural Information Processing Systems, 36.

Q&A – CGENN

- LayerNorm

$$x^{(m)} \mapsto \frac{x^{(m)}}{\sigma(a_m)\,\left(\bar{\mathfrak{q}}(x^{(m)}) - 1\right) + 1},$$

$$\langle y_i \rangle_r = \frac{\langle x_i \rangle_r}{\sigma(a_{ir})(\sqrt{\langle \langle x_i \rangle_r \langle \tilde{x}_i \rangle_r \rangle_0} - 1) + 1}$$

Ref: Ruhe, D., Brandstetter, J., and Forré, P. (2024). Clifford group equivariant neural networks. Advances in Neural Information Processing Systems, 36.