Lie Neurons: Adjoint Equivariant Neural Networks for Semi -simple Lie Algebras

UNIVERSITY OF MICHIGAN

Tzu-Yuan Lin*, Minghan Zhu*, and Maani Ghaffari

University of Michigan

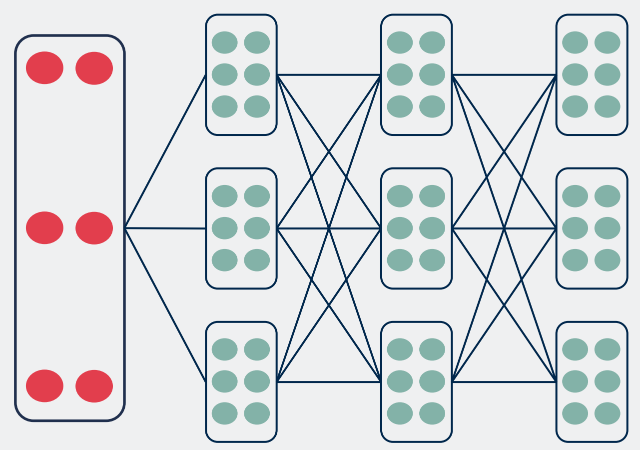
* Equal Contribution

Scan for code!

TL;DR: An MLP framework that takes Lie algebraic data as inputs and is equivariant to the adjoint representation of the group by construction.

A Lie Algebraic Network

Equivariant to adjoint actions!



Each neurons is an element in the Lie algebra

$$f(gXg^{-1}) = gf(X)g^{-1}$$

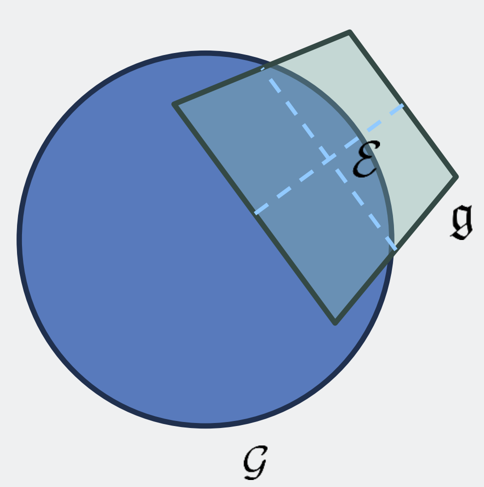
Generalize Vector Neurons (Deng et al., 2021) from $SO(n)$ to any finite dimensional semi-simple Lie algebra

Preliminaries

Lie Group & Lie Algebra

Lie group acts naturally on the Lie algebra via the **adjoint representation**.

$$Ad_g(X) = gXg^{-1}$$



Killing Form

$$B(X,Y): \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}, \quad (X,Y) \mapsto Tr(ad_X \circ ad_Y)$$

Invariant to the adjoint action

Lie Bracket

$$[X,Y] = ad_X(Y)$$
$$= XY - YX \quad \text{(For matrix Lie algebras)}$$

Equivariant to the adjoint action

Two Novel Nonlinearities

Linear

$$xW$$

Generalized ReLU

$$\begin{cases} x, & \text{if } B(x, xU) \leq 0 \\ x + B(x, xU)xU, & \text{otherwise.} \end{cases}$$

Lie Bracket

$$x + [xU, xV]$$

Geometric Channel Mixing

Improve expressiveness of the network

$$Mx$$
$$M = x_1 x_2^\intercal$$
$$x_1, x_2 = f_{\text{LN-ReLU}}(xW_i)$$

Pooling Layer

Reduce feature dimension

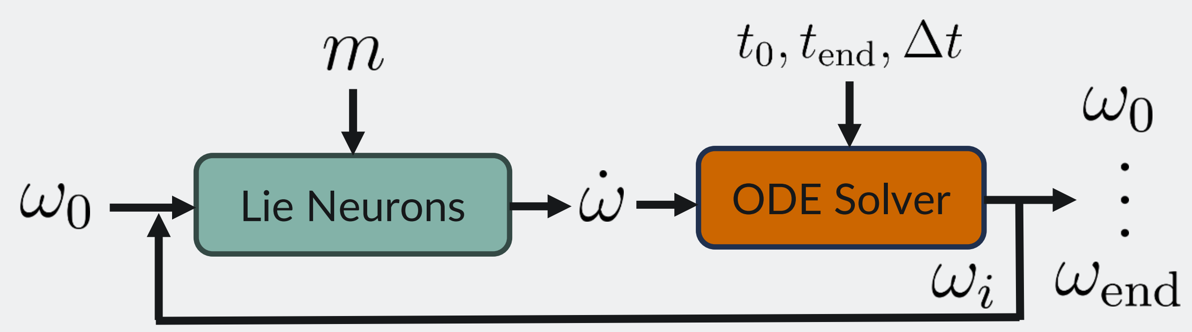
$$\arg\max_n \ B(x_n^c W, x_n^c)$$

Invariant Layer

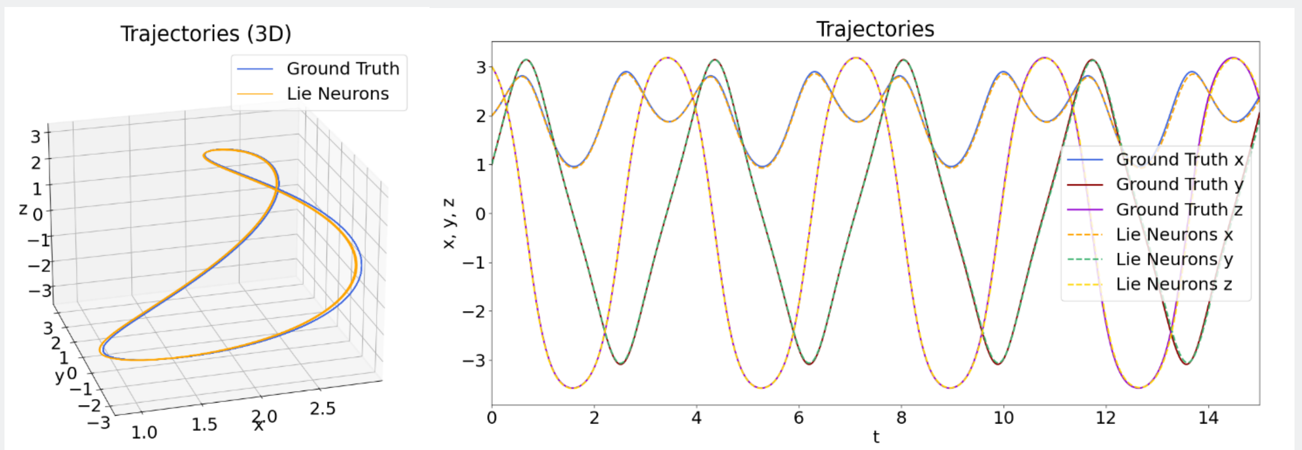
Obtain invariant features

$$B(x,x)$$

Dynamic Modeling



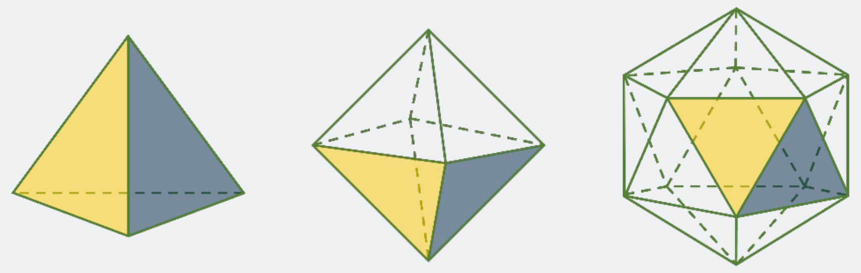
$$I\dot{\omega}(t) + \omega(t) \times I\omega(t) = 0$$



Estimation Error	Id					SO(3)				
Time (Sec)	5	10	15	20	25	5	10	15	20	25
MLP	0.428	0.656	0.717	0.763	0.799	0.474	0.689	0.733	0.768	0.805
EMLP (Finzi et al., 2021)	0.429	0.642	0.775	0.909	1.027	0.415	0.633	0.771	0.907	1.025
Lie Neurons (No Mixing)	0.739	0.842	0.791	0.805	0.809	0.739	0.842	0.791	0.805	0.809
Lie Neurons	**0.005**	**0.011**	**0.014**	**0.016**	**0.018**	**0.005**	**0.011**	**0.014**	**0.016**	**0.018**

Unit: rad/s

Platonic Solid Classification



Input: $\mathfrak{sl}(3)$ transformation between projected faces

Output: Platonic solid class

	Num Params	Accuracy		Accuracy (Rotated)	
		AVG	STD	AVG	STD
MLP	206,339	95.76%	0.65%	36.54%	0.99%
MLP Aug	206,339	81.47%	0.77%	81.20%	2.34%
LN-LR	134,664	99.56%	0.23%	99.51%	0.28%
LN-LB	200,200	99.14%	0.21%	98.78%	0.49%
LN-LR + LN-LB	331,272	**99.62%**	0.25%	**99.61%**	0.14%

ICML International Conference On Machine Learning