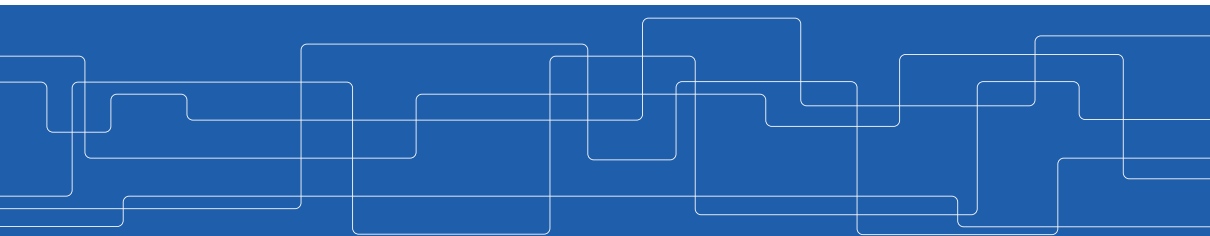




# Are All Linear Regions Created Equal?

M. Gamba, A. Chmielewski-Anders, J. Sullivan, H. Azizpour, M. Björkman

KTH Royal Institute of Technology  
Stockholm, Sweden





## Research question

- ▶ ReLU networks express continuous piece-wise affine functions of their input variable  $\mathbf{x}$ .
- ▶ To capture expressivity, interpreted as nonlinearity, estimating the number of linear “pieces” – i.e. the *density* of linear regions – has been proposed.
- ▶ However, merely *counting* linear regions heavily assumes that they all meaningfully contribute nonlinearity.
- ▶ Is this the case for standard overparameterized models?

So far, studies of linear regions were limited to small networks and MLPs.

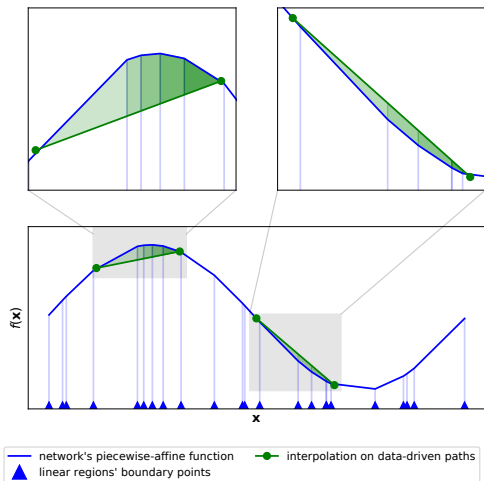
1. We propose a fast linear region discovery algorithm, that scales to realistic networks.
2. We introduce a principled measure on nonlinearity for ReLU networks.
3. We study ConvNets and ResNets trained in practice on the CIFAR datasets.
4. We show that density is an unreliable predictor in practice.
5. Our measure of nonlinearity follows the test error in a model-wise deep double descent regime.

- 
- The diagram shows a 2D space partitioned into regions by lines connecting points. Two points,  $x_0$  and  $x_n$ , are marked on the boundary. A horizontal line segment below the diagram is labeled with  $\lambda_1 d$  and  $d$ , indicating distances.

-

# Measuring Expressivity

- ▶ Study variation of the function expressed by the network.
- ▶ Measure of how far the **network** deviates from a simple **affine function**  $\mathbf{a}(\mathbf{x})$ .
- ▶  $\mathbf{a}(\mathbf{x})$  interpolates the network's function at different training points / augmentations.



- 
- The figure consists of two vertically stacked plots sharing a common x-axis representing ResNet base width, with values 12, 4, 8, 16, 24, 32, and 64. A vertical dashed line is drawn at base width 16 in both plots.
- The top plot has two y-axes. The left y-axis is 'Absolute deviation' ranging from 0 to  $1e12$ . The right y-axis is 'Density' ranging from 0 to  $1e5$ . The 'abs deviation' (solid blue line) starts at approximately  $0.2 \times 10^{12}$  at width 12, rises to a peak of about  $6.5 \times 10^{12}$  at width 16, and then decreases to about  $3.0 \times 10^{12}$  at width 64. The 'density' (dashed blue line) starts at approximately  $0.5 \times 10^5$  at width 12, increases steadily to about  $6.5 \times 10^5$  at width 64. Shaded regions around the lines indicate confidence intervals.
- The bottom plot shows 'Error' on the y-axis, ranging from 0.0 to 0.6. The 'train error' (solid blue line) starts at approximately 0.65 at width 12, drops sharply to about 0.2 at width 8, and then reaches near zero by width 16, remaining there for larger widths. The 'test error' (solid orange line) starts at approximately 0.6 at width 12, drops to about 0.35 at width 4, rises to a local maximum of about 0.5 at width 16, and then gradually decreases to about 0.38 at width 64.



Matteo Gamba



Adrian Chmielewski-Anders



Josephine Sullivan



Hossein Azizpour



Mårten Björmkän

## Code



<https://bit.ly/linear-regions>